

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:57:06 ; Search time 13.9441 Seconds
(without alignments)
2740.961 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDMQTFKA.....IAGDACLQRKNVDVSSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2599	100.0	512	4	US-09-451-739H-16
2	2578	99.2	1002	4	US-09-620-405B-475
3	2578	99.2	1002	4	US-09-604-287A-475
4	2578	99.2	1002	4	US-09-834-759-475
5	2578	99.2	1002	4	US-09-590-751A-475
6	2578	99.2	1002	4	US-09-551-621-475
7	2578	99.2	1095	4	US-09-620-405B-493
8	2578	99.2	1095	4	US-09-834-759-493
9	2125	81.8	445	4	US-09-620-405B-473
10	2125	81.8	445	4	US-09-433-826B-473
11	2125	81.8	445	4	US-09-604-287A-473
12	2125	81.8	445	4	US-09-834-759-473
13	2125	81.8	445	4	US-09-590-751A-473
14	2125	81.8	445	4	US-09-551-621-473
15	1513.5	58.2	650	4	US-09-620-405B-469
16	1513.5	58.2	650	4	US-09-433-826B-469
17	1513.5	58.2	650	4	US-09-604-287A-469
18	1513.5	58.2	650	4	US-09-834-759-469
19	1513.5	58.2	650	4	US-09-590-751A-469
20	1513.5	58.2	650	4	US-09-551-621-469
21	1513.5	58.2	743	4	US-09-620-405B-494
22	1513.5	58.2	743	4	US-09-834-759-494
23	1399	53.8	432	4	US-09-389-681-181
24	1399	53.8	432	4	US-09-620-405B-181
25	1399	53.8	432	4	US-09-339-338-181
26	1399	53.8	432	4	US-09-433-826B-181
27	1399	53.8	432	4	US-09-604-287A-181

28	1399	53.8	432	4	US-09-285-480-181	Sequence 181, App
29	1399	53.8	432	4	US-09-834-759-181	Sequence 181, App
30	1399	53.8	432	4	US-09-590-751A-181	Sequence 181, App
31	1399	53.8	432	4	US-09-551-621-181	Sequence 181, App
32	1021.5	39.3	466	4	US-09-620-405B-472	Sequence 472, App
33	1021.5	39.3	466	4	US-09-433-826B-472	Sequence 472, App
34	1021.5	39.3	466	4	US-09-604-287A-472	Sequence 472, App
35	1021.5	39.3	466	4	US-09-834-759-472	Sequence 472, App
36	1021.5	39.3	466	4	US-09-590-751A-472	Sequence 472, App
37	1021.5	39.3	466	4	US-09-551-621-472	Sequence 472, App
38	821.5	31.6	179	4	US-09-389-681-179	Sequence 179, App
39	821.5	31.6	179	4	US-09-620-405B-179	Sequence 179, App
40	821.5	31.6	179	4	US-09-339-338-179	Sequence 179, App
41	821.5	31.6	179	4	US-09-433-826B-179	Sequence 179, App
42	821.5	31.6	179	4	US-09-604-287A-179	Sequence 179, App
43	821.5	31.6	179	4	US-09-285-480-179	Sequence 179, App
44	821.5	31.6	179	4	US-09-834-759-179	Sequence 179, App
45	821.5	31.6	179	4	US-09-590-751A-179	Sequence 179, App

ALIGNMENTS

RESULT 1
US-09-451-739H-16
; Sequence 16, Application US/09451739H
; Patent No. 6774226
; GENERAL INFORMATION:
; APPLICANT: Jager, Dirk
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gure, Ali
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd
; APPLICANT: Chen, Yao-tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigen
; FILE REFERENCE: LUD 5615
; CURRENT APPLICATION NUMBER: US/09/451,739H
; CURRENT FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 16
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-451-739H-16

Query Match	100.0%	Score	2599	DB	4	Length	512
Best Local Similarity	100.0%	Pred. No.	1.7e-179				
Matches	512	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Oy	1	MKVSIPTKALELMDMQTFKAEPPEKPSAFEP	1	ESKQKDYEESSWDSLSCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120		
Db	1	MKVSIPTKALELMDMQTFKAEPPEKPSAFEP	1	ESKQKDYEESSWDSLSCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120		
Oy	61	ESKQKDYEESSWDSLSCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	61	ESKQKDYEESSWDSLSCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120		
Db	61	ESKQKDYEESSWDSLSCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	61	ESKQKDYEESSWDSLSCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM	120		
Oy	121	KVSIPTKALELMDMQTFKAEPPEKPSAFEP	121	KVSIPTKALELMDMQTFKAEPPEKPSAFEP	180		
Db	121	KVSIPTKALELMDMQTFKAEPPEKPSAFEP	121	KVSIPTKALELMDMQTFKAEPPEKPSAFEP	180		
Oy	181	SKQKVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEEDSTSLSKLDTVHSC	181	SKQKVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEEDSTSLSKLDTVHSC	240		
Db	181	SKQKVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEEDSTSLSKLDTVHSC	181	SKQKVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEEDSTSLSKLDTVHSC	240		
Oy	241	RARELQDKHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOVKVKEQLCSVRLTLN	241	RARELQDKHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOVKVKEQLCSVRLTLN	300		
Db	241	RARELQDKHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOVKVKEQLCSVRLTLN	241	RARELQDKHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOVKVKEQLCSVRLTLN	300		
Oy	301	QEEKRRNADTLNKKIREELGRIEQHKKLEVEQKQLEALRIQDIELKSVESNUNQVSH	301	QEEKRRNADTLNKKIREELGRIEQHKKLEVEQKQLEALRIQDIELKSVESNUNQVSH	360		
Db	301	QEEKRRNADTLNKKIREELGRIEQHKKLEVEQKQLEALRIQDIELKSVESNUNQVSH	301	QEEKRRNADTLNKKIREELGRIEQHKKLEVEQKQLEALRIQDIELKSVESNUNQVSH	360		

Db 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLQVSH 360
 Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
 Db 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
 Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILAEIESHHPRLASAVQDHDQI 480
 Db 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILAEIESHHPRLASAVQDHDQI 480
 Qy 481 VTSRKSQEPAPFHIAGDACLQRMNVDSSTDI 512
 Db 481 VTSRKSQEPAPFHIAGDACLQRMNVDSSTDI 512
 RESULT 2
 US-09-620-405B-475
 ; Sequence 475, Application US/09620405B
 ; Patent No. 6528054
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Hepler, William T.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.470C8
 ; CURRENT APPLICATION NUMBER: US/09/620,405B
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 495
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 475
 ; LENGTH: 1002
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(1002)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-620-405B-475
 Query Match 99.2%; Score 2578; DB 4; Length 1002;
 Best Local Similarity 99.6%; Pred. No. 1.3e-177;
 Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 60
 Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 397
 Qy 61 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
 Db 398 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
 Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 180
 Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 517
 Qy 181 SKQKXVEENSWSLSLCTVSKQDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSC 240
 Db 518 SKQKXVEENSWSLSLCTVSKQDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSC 577
 Qy 241 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWQELCSVRLTLN 300
 Db 578 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWQELCSVRLTLN 637
 Qy 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLQVSH 360
 Db 638 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLQVSH 697
 Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
 Db 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420

Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 757
 Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILAEIESHHPRLASAVQDHDQI 480
 Db 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILAEIESHHPRLASAVQDHDQI 817
 Qy 481 VTSRKSQEPAPFHIAGDACLQRMNVDSST 510
 Db 818 VTSRKSQEPAPFHIAGDACLQRMNVDSST 847
 RESULT 3
 US-09-604-287A-475
 ; Sequence 475, Application US/09604287A
 ; Patent No. 6586572
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Hepler, William T.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.470C7
 ; CURRENT APPLICATION NUMBER: US/09/604,287A
 ; CURRENT FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 489
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 475
 ; LENGTH: 1002
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: VARIANT
 ; LOCATION: (1)...(1002)
 ; OTHER INFORMATION: Xaa = Any Amino Acid
 US-09-604-287A-475
 Query Match 99.2%; Score 2578; DB 4; Length 1002;
 Best Local Similarity 99.6%; Pred. No. 1.3e-177;
 Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 60
 Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 397
 Qy 61 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
 Db 398 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
 Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 180
 Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADEILPS 517
 Qy 181 SKQKXVEENSWSLSLCTVSKQDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSC 240
 Db 518 SKQKXVEENSWSLSLCTVSKQDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSC 577
 Qy 241 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWQELCSVRLTLN 300
 Db 578 RARELOKDHCEORTGMEQMKKFCVLKKLSEAKEIKSQLENOKVWQELCSVRLTLN 637
 Qy 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLQVSH 360
 Db 638 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLQVSH 697
 Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420
 Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 757
 Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILAEIESHHPRLASAVQDHDQI 480

Db 758 KESLTKRASQYSQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 817
Qy 481 VTSRKSQEPAFHAGDACLQRMNVDVSST 510
Db 818 VTSRKSQEPAFHAGDACLQRMNVDVSST 847

RESULT 4
US-09-834-759-475
; Sequence 475, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 180
Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 517
Qy 181 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDVHSCE 240
Db 518 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDVHSCE 577
Qy 241 RARELQDHCHEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
Db 578 RARELQDHCHEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637
Qy 301 QEEKRRNADILNKEIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 360
Db 638 QEEKRRNADILNKEIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 697
Qy 361 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757
Qy 421 KESLTKRASQYSQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 758 KESLTKRASQYSQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 817
Qy 481 VTSRKSQEPAFHAGDACLQRMNVDVSST 510

Db 818 VTSRKSQEPAFHAGDACLQRMNVDVSST 847

RESULT 5
US-09-590-751A-475
; Sequence 475, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-590-751A-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 180
Db 458 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 517
Qy 181 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDVHSCE 240
Db 518 SKQKVVENSWSLSLRETYSQKDVCPKATHQKEMDKISGKLEDSTSLSKILDVHSCE 577
Qy 241 RARELQDHCHEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
Db 578 RARELQDHCHEORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637
Qy 301 QEEKRRNADILNKEIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 360
Db 638 QEEKRRNADILNKEIREELGRIBEQHRKELEVQQLQEQALRIQDIELKSVESNLNQVSH 697
Qy 361 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757
Qy 421 KESLTKRASQYSQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 758 KESLTKRASQYSQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHDQI 817
Qy 481 VTSRKSQEPAFHAGDACLQRMNVDVSST 510
Db 818 VTSRKSQEPAFHAGDACLQRMNVDVSST 847

RESULT 6

US-09-551-621-475
; Sequence 475, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; CURRENT APPLICATION NUMBER: US/09/551.621
; FILE REFERENCE: 210121.470C5
; CURRENT FILING DATE: 2000-04-17
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSTPTKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADILPS 60
Db 338 MKVSTPTKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADILPS 397

Qy 61 ESQKDYESSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESQKDYESSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 457

Qy 121 KVSITPKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADQMPFSE 180
Db 458 KVSITPKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADQMPFSE 517

Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 240
Db 518 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 577

Qy 241 RARELQKDHCEORTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 300
Db 578 RARELQKDHCEORTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 637

Qy 301 QEEKERNADILNEKIREELGRIEEQRKLELVKQQLQALRIQDIELKSVEENLNQVSH 360
Db 638 QEEKERNADILNEKIREELGRIEEQRKLELVKQQLQALRIQDIELKSVEENLNQVSH 697

Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKL 757

Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 817

Qy 481 VTSRKSQEPAFHAGDACLQKKNVDVSSST 510
Db 818 VTSRKSQEPAFHAGDACLQKKNVDVSSST 847

RESULT 7
US-09-620-405B-493
; Sequence 493, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu

; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620.405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-493

Query Match 99.2%; Score 2578; DB 4; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.4e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADILPS 60
Db 431 MKVSIPTKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADILPS 490

Qy 61 ESQKDYESSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 120
Db 491 ESQKDYESSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 550

Qy 121 KVSITPKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADQMPFSE 180
Db 551 KVSITPKALELMDMTFKAEPPKPSAFPAIEIMQKSVNPKALELKNQOTLRADQMPFSE 610

Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 240
Db 611 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKIDKINGKLEESPDNDGFLKAPCRM 670

Qy 241 RARELQKDHCEORTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 300
Db 671 RARELQKDHCEORTGMEQMKKFCVLKKLSEAKIISQLENQKVKWEQELCSVRLTLN 730

Qy 301 QEEKERNADILNEKIREELGRIEEQRKLELVKQQLQALRIQDIELKSVEENLNQVSH 360
Db 731 QEEKERNADILNEKIREELGRIEEQRKLELVKQQLQALRIQDIELKSVEENLNQVSH 790

Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKL 420
Db 791 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKL 850

Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 851 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 910

Qy 481 VTSRKSQEPAFHAGDACLQKKNVDVSSST 510
Db 911 VTSRKSQEPAFHAGDACLQKKNVDVSSST 940

RESULT 8
US-09-834-759-493
; Sequence 493, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.

APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C9
CURRENT APPLICATION NUMBER: US/09/834,759
CURRENT FILING DATE: 2001-04-13
NUMBER OF SEQ ID NOS: 547
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 493
LENGTH: 1095
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(1095)
OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493

Query Match 99.2%; Score 2578; DB 4; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.4e-177;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 60
Db 431 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 490
Qy 61 ESKQDYESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 491 ESKQDYESSWDSSESLCETVSKQDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRM 550
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 180
Db 551 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 610
Qy 181 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGLKEDSTSLSKILDTVHSC 240
Db 611 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGLKEDSTSLSKILDTVHSC 670
Qy 241 RARELQDHCEQRTGKQMKKFCVLKCKLSEAKEIKSQLENQKVWEQELCSVRLTLN 300
Db 671 RARELQDHCEQRTGKQMKKFCVLKCKLSEAKEIKSQLENQKVWEQELCSVRLTLN 730
Qy 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 731 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 790
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTKL 420
Db 791 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTKL 850
Qy 421 KEESLTKRASQYSGQLKVLTAENTMLTSKLKQKODKEILBAEIESHPRLASAVQDDHDI 480
Db 851 KEESLTKRASQYSGQLKVLTAENTMLTSKLKQKODKEILBAEIESHPRLASAVQDDHDI 910
Qy 481 VTSRKSQEPAPHIAGDAQLQRMNVDSST 510
Db 911 VTSRKSQEPAPHIAGDAQLQRMNVDSST 940

RESULT 9
US-09-620-405B-473
Sequence 473, Application US/09620405B
Patent No. 6528054
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqui
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

FILE REFERENCE: 210121.470C8
CURRENT APPLICATION NUMBER: US/09/620.405B
CURRENT FILING DATE: 2000-07-20
NUMBER OF SEQ ID NOS: 495
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 473
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-09-620-405B-473
Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 60
Db 26 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 85
Qy 61 ESKQDYESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSSESLCETVSKQDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 180
Db 146 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADMPFSE 205
Qy 181 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGLKEDSTSLSKILDTVHSC 240
Db 206 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGLKEDSTSLSKILDTVHSC 265
Qy 241 RARELQDHCEQRTGKQMKKFCVLKCKLSEAKEIKSQLENQKVWEQELCSVRLTLN 300
Db 266 RARELQDHCEQRTGKQMKKFCVLKCKLSEAKEIKSQLENQKVWEQELCSVRLTLN 325
Qy 301 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIBEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 385
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMT 417
Db 386 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMT 442

RESULT 10
US-09-433-826B-473
Sequence 473, Application US/09433826B
Patent No. 6579973
GENERAL INFORMATION:
APPLICANT: Jiang, Yuqui
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.470C4
CURRENT APPLICATION NUMBER: US/09/433.826B
CURRENT FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 474
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 473
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-09-433-826B-473

Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEWQKSVPNKALELKNQETLRADIELPS 60

Db 26 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADEILPS 85
Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSLSLCTVTSQKDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADOMFPE 180
Db 146 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADOMFPE 205
Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 240
Db 206 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 265
Qy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOKVWQEQALRIQDIELKSVEENLNQVSH 300
Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOKVWQEQALRIQDIELKSVEENLNQVSH 325
Qy 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

RESULT 11
US-09-604-287A-473
; Sequence 473, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-604-287A-473

Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADEILPS 60
Db 26 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADEILPS 85
Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSLSLCTVTSQKDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADOMFPE 180
Db 146 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADOMFPE 205
Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 240
Db 206 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 265
Qy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOKVWQEQALRIQDIELKSVEENLNQVSH 300
Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOKVWQEQALRIQDIELKSVEENLNQVSH 325
Qy 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOKVWQEQALRIQDIELKSVEENLNQVSH 325
Qy 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

RESULT 12
US-09-834-759-473
; Sequence 473, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-473

Query Match 81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADEILPS 60
Db 26 MKVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADEILPS 85
Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYESSWDSLSLCTVTSQKDVCLPKAAHQKEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADOMFPE 180
Db 146 KVSIPKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELKNQTLRADOMFPE 205
Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 240
Db 206 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCE 265
Qy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOKVWQEQALRIQDIELKSVEENLNQVSH 300
Db 266 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENOKVWQEQALRIQDIELKSVEENLNQVSH 325
Qy 301 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIBEOHREKLEVKQEQALRIQDIELKSVEENLNQVSH 385
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 417
Db 386 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMT 442

RESULT 13
US-09-590-751A-473
; Sequence 473, Application US/09590751A
; Patent No. 6756477

```

; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-590-751A-473

Query Match      81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRADEILPS 60
Db 26 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRADEILPS 85
Qy 61 ESKQDYEESSWDSSESLCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYEESSWDSSESLCETVSQKDVCLPKAAHQEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRAQDMPFSE 180
Db 146 KVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRAQDMPFSE 205
Qy 181 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLSKILDTVHSC 240
Db 206 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLSKILDTVHSC 265
Qy 241 RARELQDHCERTGKMEQMKKFCVLKKLSEAKETKSOLENKVQWQELCSVRLTLN 300
Db 266 RARELQDHCERTGKMEQMKKFCVLKKLSEAKETKSOLENKVQWQELCSVRLTLN 325
Qy 301 QEEKRRNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 385
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMT 417
Db 386 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMT 442

RESULT 15
US-09-620-405B-469
; Sequence 469, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-469

Query Match      58.2%; Score 1513.5; DB 4; Length 650;
Best Local Similarity 99.0%; Pred. No. 4e-101;
Matches 297; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRADEILPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRADEILPS 397
Qy 61 ESKQDYEESSWDSSESLCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120

```

```

; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-551-621-473

Query Match      81.8%; Score 2125; DB 4; Length 445;
Best Local Similarity 99.8%; Pred. No. 2e-145;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRADEILPS 60
Db 26 MKVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRADEILPS 85
Qy 61 ESKQDYEESSWDSSESLCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 86 ESKQDYEESSWDSSESLCETVSQKDVCLPKAAHQEIDKINGKLEESPDNDGFLKAPCRM 145
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRAQDMPFSE 180
Db 146 KVSIPTKALELMDMOTFKAEPPKPSAFPAIEMQKSVPNKALELNEQTLRAQDMPFSE 205
Qy 181 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLSKILDTVHSC 240
Db 206 SKQKVEENSWDSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTSLSKILDTVHSC 265
Qy 241 RARELQDHCERTGKMEQMKKFCVLKKLSEAKETKSOLENKVQWQELCSVRLTLN 300
Db 266 RARELQDHCERTGKMEQMKKFCVLKKLSEAKETKSOLENKVQWQELCSVRLTLN 325
Qy 301 QEEKRRNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 326 QEEKRRNADILNEKIREELGRIEEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSH 385
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMT 417
Db 386 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMT 442

RESULT 14
US-09-551-621-473
; Sequence 473, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-551-621-473

```

Db 398 ESKQXDYESSWDSLSCTVSQKDVCLPKAXHQKEIDKINGKLBESPNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALTEMOKSVPNKALELKNBQTLRADQMPFSE 180
Db 458 KVSIPTKALELMDMOTFKAEPPKPSAFEPALTEMOKSVPNKALELKNBQTLRADQMPFSE 517
Qy 181 SKQKXVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCE 240
Db 518 SKQKXVEENSWDSLSRETYSQKDVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCE 577
Qy 241 RARELQKHCEORTGMEOMKKKFCVLKKLSEAKEIKSLENQKVKWEQELCSVR-LTL 299
Db 578 RARELQKHCEORTGMEOMKKKFCVLKKLSEAKEIKSLENQKVKWEQELCSVRFLT 637

Search completed: February 19, 2005, 02:13:26
Job time : 14.9441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:55:34 ; Search time 62.9274 Seconds
(without alignments)
4166.462 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDMQTFKA.....TAGDACLQRKXNDVSSSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2589	99.6	1341	2 Q9BXX3	Q9bxx3 homo sapien
2	1542	59.3	1011	2 Q9BXX2	Q9bxx2 homo sapien
3	1094	42.1	424	2 Q9NS19	Q9ns19 homo sapien
4	927.5	35.7	1715	2 Q9UP88	Q9up88 homo sapien
5	927	35.7	1710	2 Q9HIQ1	Q9hiq1 homo sapien
6	871	33.5	1080	2 Q6ZRI4	Q6zri4 homo sapien
7	650.5	25.0	718	2 Q9HCD1	Q9hcd1 homo sapien
8	612.5	23.6	1043	2 Q69ZS2	Q69zs2 mus musculu
9	519	20.0	823	2 Q9H0H6	Q9h0h6 homo sapien
10	475	18.3	453	2 Q81ZM7	Q81zm7 homo sapien
11	431.5	16.6	641	2 Q60311	Q60311 homo sapien
12	313	12.0	2006	2 Q7K5Q6	Q7k5q6 plasmodium
13	313	12.0	2019	2 Q7K5Q5	Q7k5q5 plasmodium
14	313	12.0	2055	2 Q8IHP3	Q8ihp3 plasmodium
15	313	12.0	2055	2 Q8T5C7	Q8t5c7 plasmodium
16	294.5	11.3	520	2 Q80W27	Q80w27 mus musculu
17	294.5	11.3	992	1 AN18_HUMAN	Q8ivf6 homo sapien
18	292.5	11.3	381	2 Q15694	Q15694 homo sapien
19	267.5	10.3	292	2 Q68DM0	Q68dm0 homo sapien
20	265	10.2	2954	2 Q42263	Q42263 xenopus lae
21	264	10.2	450	2 Q6PI49	Q6pi49 homo sapien
22	264	10.2	1738	2 Q76329	Q76329 dictyosteli
23	261.5	10.1	2116	1 MYS2_DICTDI	P08799 dictyosteli
24	255	9.8	2057	1 MYSN_DROME	Q99323 drosophila
25	251.5	9.7	1305	2 Q9FJ35	Q9fj35 arabidopsis
26	250	9.6	944	1 NUFI1_YEAST	P32380 saccharomyc
27	247.5	9.5	2230	1 GOA4_HUMAN	Q13439 homo sapien
28	247.5	9.5	2253	2 P70012	P70012 xenopus lae
29	247	9.5	1993	2 Q7PF91	Q7pf91 anopheles g
30	247	9.5	2016	2 Q7PF90	Q7pf90 anopheles g
31	245.5	9.4	1762	2 Q94DC2	Q94dc2 oryza sativ

32 244.5 9.4 1979 2 Q96133 096133 plasmodium
33 244 9.4 1980 2 Q6FWE0 Q6fwe0 candida gla
34 244 9.4 7210 2 Q9V7G8 Q9v7g8 drosophila
35 244 9.4 9270 2 Q8MLD9 Q8mld9 drosophila
36 242.5 9.3 689 2 Q80VF0 Q80vf0 mus musculu
37 242.5 9.3 692 2 Q9DSR3 Q9dsr3 mus musculu
38 241.5 9.3 1313 2 Q9XIP6 Q9xip6 arabidopsis
39 241.5 9.3 2042 2 Q6BUO9 Q6buq9 debaryomyce
40 240.5 9.3 997 1 SCPI1_RAT Q03410 rattus norv
41 240.5 9.3 1320 2 Q9JKZ5 Q9jkz5 rattus norv
42 239.5 9.2 1671 2 Q8MQ60 Q8mq60 caenorhabdi
43 239.5 9.2 1827 2 Q20042 Q20042 caenorhabdi
44 239.5 9.2 1885 2 Q95QG9 Q95qg9 caenorhabdi
45 239.5 9.2 1898 2 Q95QH0 Q95qh0 caenorhabdi

ALIGNMENTS

RESULT 1

Q9BXX3 ID Q9BXX3 PRELIMINARY; PRT; 1341 AA.
AC Q9BXX3, 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DB Breast cancer antigen NY-BR-1.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL: AF269087; AAK27325.1;
DR GO: GO:0005634; C:nucleus; NAS.
DR GO: GO:0005515; F:protein binding; NAS.
DR GO: GO:0003700; F:transcription factor activity; NAS.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR001969; Pept_Asp_AS.
DR Pfam: PF00023; Ank; 6.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 6.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00141; ASP_PROTEASE; UNKNOWN_1.
KW ANK repeat.
SQ SEQUENCE 1341 AA; 152776 MW; 33E53DDE6FD3A58B CRC64;
Query Match 99.6%; Score 2589; DB 2; Length 1341;
Best Local Similarity 100.0%; Pred. No. 2.2e-96;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMQTFKAPPEKPSAFPAIEQKSVPNKALELKNQETLRADEILPS 60
Db 677 MKVSIPTKALELMDMQTFKAPPEKPSAFPAIEQKSVPNKALELKNQETLRADEILPS 736
Qy 61 ESKQKDYESSWDSSESICETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQKDYESSWDSSESICETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
Qy 121 KVSITPTKALELMDMQTFKAPPEKPSAFPAIEQKSVPNKALELKNQETLRAQMPFSE 180
Db 797 KVSITPTKALELMDMQTFKAPPEKPSAFPAIEQKSVPNKALELKNQETLRAQMPFSE 856
Qy 181 SKQKVKVENSWSLSRETYSQKDVCPKATHQKEMDKISKLEBDSLSKILDTVHSCB 240

```

Db      857 SKQKVEENSWSBSLRTVSQKVCVPKATHQKEMDKISGLKEDSTLSKILDTVHSCE 916
Qy      241 RARELQKDHCBORTGKMEQMKKFCVLKKLSEAKBIKSOLENQKVKEQELCSVRLTLN 300
Db      917 RARELQKDHCBORTGKMEQMKKFCVLKKLSEAKBIKSOLENQKVKEQELCSVRLTLN 976
Qy      301 QEEKRRNRADILNEKIRELGRIEHQHREKLEVKQLEQALRIODIELKSVESNLNOVSH 360
Db      977 QEEKRRNRADILNEKIRELGRIEHQHREKLEVKQLEQALRIODIELKSVESNLNOVSH 1036
Qy      361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db      1037 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy      421 KEESLTKRASQVSGQKVLIAENTMTLSKLKEKQDKTEILEAEIESHHPRILASAVQDHDQI 480
Db      1097 KEESLTKRASQVSGQKVLIAENTMTLSKLKEKQDKTEILEAEIESHHPRILASAVQDHDQI 1156
Qy      481 VTSRKSQBPAPHIAGDACLQRKNVDSST 510
Db      1157 VTSRKSQBPAPHIAGDACLQRKNVDSST 1186

RESULT 2
Q9BXK2 ID Q9BXK2 PRELIMINARY; PRT; 1011 AA.
AC Q9BXK2; MEDLINE=21174979; PubMed=11280766;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert B., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269088; AK27326.1; -.
DR HSSP; O75832; IUOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
FT NON_TER 1011 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DE438 CRC64;

Query Match 59.3%; Score 1542; DB 2; Length 1011;
Best Local Similarity 50.5%; Pred. No. 2e-54;
Matches 336; Conservative 42; Mismatches 68; Indels 220; Gaps 5;

Qy      2 KVSIPTKALE-LMDMOTFKAEPPKPSAFEPAIEMQKSVNPKALELKNQOTLRADIELPS 60
Db      351 QVCIPESMTQKWEINREVEELPEKPSAFKPAVEMQKTVNKAFLKNEQTLRAAQMPFS 410
Qy      61 ESKQKDYESSWSDESLSCTVSQKQVCLPKATHQKIDKINGKLRESPDNDGFLKAPCRM 120
Db      411 ESKQKDYESSWSDESLSCTVSQKQVCLPKATHQKIDKINGKLRESPDNDGFLKAPCRM 470
Qy      121 KVSIPTKALELMDMOTFKAEPPKPSAFEP-----AIEMQK----- 156
Db      471 KVSIPTKALELMDMOTFKAEPPKPSAFEPDKGLLKPTCGKVSIPNPKALELKDRETLKAESPNDNG 530
Qy      157 -----SVNPKALELKNQOTLRADQMFPSKQKQVENSWSDESLSRETVSQKVC 206
Db      531 LKPKTCGRKVSIPNPKALELKDRETFKAAQMPFSKQKQVENSWSDESLSRETVSQKVC 590

```

```

Qy      207 VPKATHQKEMDKISGKLE----- 224
Db      591 LPKATHQKEDFTLSGKLEESPDGGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVES 650
Qy      225 ----- 224
Db      651 TFSLFGKPTTENSQSTKVBEEDFNLTTKEGATKVTGQQRDGIIGIERAPQDQTNKMPTE 710
Qy      225 ----- 224
Db      711 LGRKEDTKTSDSEIISVSDTQNYECLPEATYQKEIKTTNGKIEESPEKSPHFEPATEMQ 770
Qy      225 -----DSTLSKILDTVHSERARELOKDHCEQRTGKMEQMKKFCV 266
Db      771 NSVPNKGLEWKNQTLRADSTTILSKILDALPSGERGELKDNCEQITANWEQMKKFCV 830
Qy      267 LKKLSEAKBIKSOLENQKVKEQELCSVRLTNQEEBEKERNADILNEKIRELGRIEHQ 326
Db      831 LKELSEAKBIKSOLENQKAKWEQELCSVRLTNQEEBEKERNVDILKEKIRP-----EEQ 885
Qy      327 HRKELEVKQLEQALRIODIELKSVESNLNOVSHTHENENYLLHNCMLKKEIAMLKLEI 386
Db      886 LRKLEVKHQLEQTLRIQDIELKSVTSNLNOVSHSENDLFHENCMLKKEIAMLKLEI 945
Qy      387 ATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQVSGQKVLIAENTML 446
Db      946 ATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKQKTVTKRASQVREQKVLTAENTML 1005
Qy      447 TSKLKE 452
Db      1006 TSKLKE 1011

RESULT 3
Q9NSI9 ID Q9NSI9 PRELIMINARY; PRT; 424 AA.
AC Q9NSI9;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE PRED4 protein (Fragment).
GN Name=PRED4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,
RA Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,
RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sakaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordisiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.L.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL163203; CAB90394.1; -.
FT NON_TER 1 424
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

Query Match 42.1%; Score 1094; DB 2; Length 424;
Best Local Similarity 79.0%; Pred. No. 8.3e-37;
Matches 226; Conservative 19; Mismatches 27; Indels 14; Gaps 2;

Qy      225 DSTLSKILDTVHSERARELOKDHCEQRTGKMEQMKKFCVVKKLESEAKBIKSOLENQ 284

```

```

Db 2 DSTLSKILDVAPSCERGKELKKDCEQITAKWEQTKNFKCVLQKELSEAKEIKSQLENO 61
Qy 285 KVKWEQELCSVRLTLNQBEERKNADILNEKIREBLGRIEHQHKEVKKQLEQALRIQ 344
Db 62 KAKWEQELCS-----KKRDVDILKEKIRP-----EEQURKKLEVKQLEQALRIQ 107
Qy 345 DIELKSVSNLNOVSHTHENENYLHENCMLKKEIAMLKLEIATILKHQYQEKENKYPEDI 404
Db 108 DIELKSVTSNLNOVSHTHESENDLPHENCMLKKEIAMLKLEIATILKHQYQEKENKYPEDI 167
Qy 405 KILKEKNAELQWTLKKEESLTKRASQYSGOLKVLIAENTMLTSLKKEKQDKEILEAEIE 464
Db 168 KILQENKNAELQWTLKQKTLTKRASQYREQKVLTAENTMLTSLKKEKQDKEILEAEIE 227
Qy 465 SHHPLASAVQDHDQIVTSRKSQEPAFHAGDACLQKQKNDVDSST 510
Db 228 SHHPLASALQDHDQSVTSRKNQELAFHAGDAHLQGIQMDVDVSN 273

RESULT 4
Q9UPS8 PRELIMINARY; PRT; 1715 AA.
AC Q9UPS8;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)
DE KIAA1074 protein (Fragment).
GN Name=KIAA1074;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirotsawa M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
DR EMBL; AB028997; BAAB3026.2; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
FT NON_TER
SQ SEQUENCE 1715 AA; 196962 MW; A82858F5F58E3203 CRC64;

Query Match 35.7%; Score 927.5; DB 2; Length 1715;
Best Local Similarity 40.7%; Pred. No. 1.6e-29;
Matches 233; Conservative 94; Mismatches 156; Indels 89; Gaps 13;

Qy 21 EPPEKPSAFEPALIEQKSVNPKALELKNQETLRADE-ILPSESKQDYESSWDSSESLCE 79
Db 494 ESPERYHLKPTIEMKDSVPNKAQGMKDVQTSKAAAEHDLVASEEEQREGSENNQPOVE 553
Qy 80 TVSQKDVCLPKATHQKEI-----DKINGKLESPNDGFL---KAPCRMKVSIPTK---- 127
Db 554 EERKK-----HRNNEVEVSANIHDGATDDAEDDDDDGLIQRKSGETHQOFPKENKE 608
Qy 128 -----ALELMDMQ-TFKAEPPKPSAFEPALIEQKSVNPKALELKNQETL-----RA 173
Db 609 YASGPALQMKVEKSTESKESVNSPVFGKASLLTGGLLQVDDSSLSIDEDEGR 668
Qy 174 DQMPFSESKQKVEENSWDS-----ESLRETVSQKDVCPKATHQKEM---DKISGLKEDS 226
Db 669 TKTKTSNEKNKVNQIQSQMDVDDLTQSSETASE-DCLELPHSSYKNFMILLIEQLGMECKDS 727

```

```

Qy 227 TSLSKILDTVHSCERARELQKHCHCQRTGQMEQMKKFCVLKKKLSAKEIKSQLENO 286
Db 728 VSLKIQDAALSCERLLSLKKNHCELLTVKIKMEDKVNVLQRELSSETKIKSQLEHQKV 787
Qy 287 KWQELCSVRLTLNQBEERKNADILNEKIREBLGRIEHQHKEVKKQLEQALRIQDI 346
Db 788 EMERELCSLRFSLNQBEERKNADILYKIREQLRKRKEEQYRKEVEVKKQLESLQTLDEM 847
Qy 347 ELKSVESENLNOV-----SH 360
Db 848 ELRTVKSNLNQVQERNDAQRLSPEQARNMLQDGLITNHLSSKQKEIEWAKKMNSENH 907
Qy 361 THENENYLHENCMLKKEIAMLKLEIATILKHQYQEKENKYPEDIKILKEKNAELQWTLK 420
Db 908 SHEEKOLSHKNMLQEBIAMLRLEIDITIKNQOEKKKCPEDLKIYKKNEDLQKTIKQ 967
Qy 421 KEESLTKRASQYSGOLKVLIAENTMLTSLK-K-EKQDKEILEAEIESHPRLASAVQDHDQ 479
Db 968 NEETLTQTISQYNGRLSVLTAEANMLSKLENKOSKERLEAEVESHSLAAATHDRDQ 1027
Qy 480 IVTSRKQEPAFHAGDAC--LQRKNVDVSS 509
Db 1028 SETSKRELELAFQARDECRLQDKMNFVSN 1059

RESULT 5
Q9HIQ1 PRELIMINARY; PRT; 1710 AA.
AC Q9HIQ1;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DE 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE BA145B8.1 (KJAA1074).
GN Name=ba145B8.1;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1710 AA; 196409 MW; 01CBF9BADB894872 CRC64;

Query Match 35.7%; Score 927; DB 2; Length 1710;
Best Local Similarity 40.7%; Pred. No. 1.7e-29;
Matches 233; Conservative 94; Mismatches 156; Indels 90; Gaps 13;

Qy 21 EPPEKPSAFEPALIEQKSVNPKALELKNQETLRADE-ILPSESKQDYESSWDSSESLCE 79
Db 488 ESPERYHLKPTIEMKDSVPNKAQGMKDVQTSKAAAEHDLVASEEEQREGSENNQPOVE 547
Qy 80 TVSQKDVCLPKATHQKEI-----DKINGKLESPNDGFL---KAPCRMKVSIPTK---- 127
Db 548 EERKK-----HRNNEVEVSANIHDGATDDAEDDDDDGLIQRKSGETHQOFPKENKE 602
Qy 128 -----ALELMDMQ-TFKAEPPKPSAFEPALIEQKSVNPKALELKNQETL-----R 172
Db 603 YASGPALQMKVEKSTESKESVNSPVFGKASLLTGGLLQVDDSSLSIDEDEGR 662
Qy 173 ADQMPFSESKQKVEENSWDS-----ESLRETVSQKDVCPKATHQKEM---DKISGLKED 225
Db 663 TKTKTSNEKNKVNQIQSQMDVDDLTQSSETASE-DCLELPHSSYKNFMILLIEQLGMECKD 721

```

```

QY 226 STSLKILDTVHSCERARELOKHCEQRTGMEQMKKFCVLKKLSEAKEIKSOLENOX 285
DB 722 SVSLKIQDAALSCEBRLLEKKNHCELLTVKIKMEDKVNVLQRELSSETKEI 781
QY 286 KWVEQELCSVRLTNOEBEKNRNADILNEKIREELGRIEEQRKELEVKQOLEQALRIQD 345
DB 782 VWEERELCSLRSFNQEBEKNRNADTLYEKIREQLRRKEEQYRKEVEVKQLELSLQTL 841
QY 346 IELKSVESNLQV-----S 359
DB 842 MELRTVKSNLNVQVQERNDQAQOLSRQEQNARMQDGLTILNHLKQKEIEMAKKMSNS 901
QY 360 HTHENENYLLHNCMLKEIATLKHQYOEKENYFEDIKILKEKNAELQMTLK 419
DB 902 HSHEEKKDLSHNSMLQEBIAMLRLIEDITIKNQOEKKCFEDUKI VKEKNEDLQKTIK 961
QY 420 LKEESITKRAQSYGOLKVLIATNMTLSKLK-EKQDKIILAEIESHPHRLASAVQDHD 478
DB 962 QNEEILTQTSYNGRLSVLTAEANMLNSKLENEKQSKERLEAEVESHSLAAIHDRD 1021
QY 479 QIVTGRKSOEPAFHAGDAC--LQRKNVDVSS 509
DB 1022 QSETSKRELEAFQARDECSRLQDKMNFVSN 1054

RESULT 6
Q6ZR14 PRELIMINARY; PRT; 1080 AA.
AC Q6ZR14;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE K1AA1641 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Trachea;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Tahii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuna M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakani B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128577; BAC87508.1; -
FT NON_TER 1080 1080
SQ SEQUENCE 1080 AA; 126741 MW; 47BBAAA933AE595C CRC64;

Query Match 33.5%; Score 871; DB 2; Length 1080;
Best Local Similarity 38.7%; Pred. No. 1.9e-27;
Matches 225; Conservative 97; Mismatches 153; Indels 106; Gaps 14;

QY 2 KVSIPKALELMDQTFKAE-----PPEKSAFPALEIMQKSVNKALELKN---EOTLRA 54
DB 64 KVSLEELQ--DMQRFKNEIGMLKVEFQALEKEVQLOKEVEERKGRHNEVESANI 121
QY 55 DEILPSESOKOYBESSWDSSELCETVSQKDVCLPK-----ATHQKEIDKINGK 103
DB 122 HDGATDAEDDDDDGLIQKSGETHHQ---FPRKENKEYASSGALQMKVEKSTEXE 178
QY 104 LEFSPNDGFLKAPCRMKVSIPTKALELMDQTFKABPP-----KPSAFPALEIMQKSVN 160
DB 179 KRTSKES---VNSPVFGKASLLTGLLIQVDDDSLSLSEIDEDEGRPT-----KKTSN 226
QY 161 KALELKNQTLRADQMPFSESQKQKVEENSWDSSELSRETVSQKDVCPKATHQKEM---D 217
DB 227 EKNKVRNQIQSMDD-----VDDLQSSSETASE-----DCELPSSYKNFMLLIE 270

```

```

QY 218 KISKLEEDSTSLKILDTVHSCERARELOKHCEQRTGMEQMKKFCVLKKLSEAKEI 277
DB 271 QLGWECKDSVSLKIQDAALSCEBRLLEKKNHCELLTVKIKMEDKVNVLQRELSSETKEI 330
QY 278 KSQLENOKVDWEQELCSVRLTNOEBEKNRNADILNEKIREELGRIEEQRKELEVKQOL 337
DB 331 KSQLEHQKVEWERELCSLRSFNQEBEKNRNADTLYEKIREQLRRKEEQYRKEVEVKQOL 390
QY 338 EQALRIODIELKSVESNLQV----- 358
DB 391 ELSLQTLMELELRTVKSNLNVQVQERNDQAQOLSRQEQNARMQDGLTILNHLKQKEIEMAQ 450
QY 359 -----SHTHENENYLLHNCMLKEIATLKHQYOEKENYFEDIKILKEKN 411
DB 451 KKNSENHSHHEBEKOLSHKNSMLQEBIAMLRLIEDITIKNQOEKKCFEDILKIVKEKN 510
QY 412 AELQMTLKLKEESITKRAQSYGOLKVLIATNMTLSKLK-EKQDKIILAEIESHPHRL 470
DB 511 EDLQKTIKQNEEILTQTSYNGRLSVLTAEANMLNSKLENEKQSKERLEAEVESHSL 570
QY 471 ASAVODHDQIVTGRKSOEPAFHAGDAC--LQRKNVDVSS 509
DB 571 AAATHDRQSETSKRELEAFQARDECSRLQDKMNFVSN 611

RESULT 7
Q9HCD1 PRELIMINARY; PRT; 718 AA.
AC Q9HCD1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE K1AA1641 protein (Fragment).
GN Name=K1AA1641;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hiroseawa M., Ohara O.;
RT Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.
RL DNA Res. 7:273-281 (2000).
DR EMBL; AB046861; BAB13467.1; -
FT NON_TER 1
SQ SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;

Query Match 25.0%; Score 650.5; DB 2; Length 718;
Best Local Similarity 34.9%; Pred. No. 9.3e-19;
Matches 177; Conservative 102; Mismatches 185; Indels 43; Gaps 12;

QY 24 EKPSAFPALEIMQKSVNKALELKNQTLRADQILPSESOKOYBESSWDSSELSRETVSQ 83
DB 13 QKQPAKATSKOSVSNIPTEIKDQGO-----SGTVSSQK---QPAWKATSV-----K 58
QY 84 KDVCPLKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPPE 143
DB 59 KDSVSNIAIEKD-QQIRGTV--SPQKSAQKVFPPKKVLLNIAITRITGGWKSGETYPE 115
QY 144 KPSAFPALEIMQKSVNKALELKNQTLRADQ--MPPSESQKQKVEENSWDSSELSRETVS 201
DB 116 NLPTLKATIENKNSVLNTATKMKDVQSTPTEQDLEMASEGEQKLEEEYENQPVKNQIH 175
QY 202 QK-----DVCVPKATHQKE-----MDKISGKLEDSLSLSKILDTVHSCERAR 243
DB 176 SRDDLDDLIQSSQTSVEDGDSLCCNCKNVILLIDHEMKCKDCVHLKIKNTFFCLMKRLI 235
QY 244 ELQDKHCEQRTGMEQMKKFCVLKKLSEAKEIKSOLENOKVKWEQELCSVRLTNOEBE 303
DB 236 KLKDNHCEQLRKIRKLNKKNKASVLQKRISKEBEIKSQLKHEILELEKELCSLRFATQBEK 295

```



```

Qy 304 EKRRNADILNEKIRELGRIBEQHRKELEVKQOQLALRIODIELKSVESNLNQVSHTHE 363
Db 296 KKRNVHELHOKVREKLITBEQVRIEADVTPIKPAKLSAEVLKGTGNNNSQVSETDE 355
Qy 364 NENYLLHNCMLKKEIAMKLEIATLKHQYOEKENYFEDIKILKEKNAELQMTLKLKEE 423
Db 356 KED-LLHNRWQDEIARLEKDTIKNQNEK--KYLKDFEIVKRRHEDIQKALKRNGE 412
Qy 424 SLTKRASQYSGQLKVLIAENTMLTSLK--KEKQDKEILEAEIESHPRLASAVQDHDQIVT 482
Db 413 TLAKTIACYSQLAALTIDENTITLTKLEKQESRQRLTEMQSYHCRNLNARCDHQSHS 472
Qy 483 SRKSQEPAFHIAGDAC--LQRKMNVDV 507
Db 473 SKRDQELAFQGTVDKCRHLQENLNSHV 499

RESULT 8
Q69ZS2
ID Q69ZS2 PRELIMINARY; PRT; 1043 AA.
AC Q69ZS2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE MKIAA1074 protein (Fragment).
GN Name=MKIAA1074;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Soga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004)
DR EMBL; AK173096; BAD32374.1; -.
DR InterPro; IPR010989; t-snare.
DR InterPro; IPR009054; Topismrse_insert.
FT NON TER 1
FT NON TER 1043
SQ SEQUENCE 1043 AA; 121354 MW; 5F37B1F9AA37B212 CRC64;

Query Match 23.6%; Score 612.5; DB 2; Length 1043;
Best Local Similarity 32.1%; Pred. No. 4.5e-17;
Matches 177; Conservative 84; Mismatches 143; Indels 147; Gaps 15;

Qy 20 AEPKESAFPEAEMQKSVPNKALELKNEQTLR-----ADEILPSESQKQDYESS 71
Db 1 AEP--SGNLYSGAAGDGVAPQSGDTEHQPSREGSGRGGPALLMKEAKGVENEK-- 56
Qy 72 WDSLSCTVTSQKDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALEL 131
Db 57 WVSREPARTANSERTGLP-----TGQWQMDGSCW----- 87
Qy 132 MDMTQTKAEPKPSAFPEAEMQKSVPNKALELKNEQTLRADQM----FPSESQKQKVE 187
Db 88 SDTQSEARPTKTS-----SKHNK---DSGQTAADVNDLDDFTSESTASEDHE 133
Qy 188 ENSWDSLERTVTSQKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSCERARELOK 247
Db 134 LQGPDSESI-----LCA-----IEHLRECKDTASLLKIRDAVYSYKSLIELKR 177
Qy 248 DHCORTGKMBQMKKFCVLKLLSEAKEIKSOLENOKVKEQELCSVRLTLNQBEKKRR 307
Db 178 SHCELLTGKLRMENKYGLQKEMSETEVSKRLEHEKVGWEQELCRLRFALKQBEKKRR 237
Qy 308 NADILNEKIRELGRIBEQHRKELEVKQOQLALRIODIELKSVESNLNQV----- 358

```

```

Db 238 SADQLSEKTMQLRRKKEQCOSEVAREQOQLASRLTLEMELKTVKSHLNQVLEERNETQR 297
Qy 359 -----SH-----THENYLLHNCMLKKEIAM 381
Db 298 QLSREQNARMLQDGLASHLCKQKEIEMTKKMTSEVSVSHKEKDLHLKHQRLQDEVAV 357
Qy 382 LKLEIATLKHQYQKBNKYFEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIA 441
Db 358 LRLEMDTIKSHNQEKRYLEDIKIANEKDNLQRMVKL----- 396
Qy 442 ENTMLTSLK--KEKQDKEILEAEIESHPRLASAVQDHDQIVTSRKSPQAFHIAGDAC-- 498
Db 397 --NNMLSSKLDNEKQKQRLTDEVSFRSLASALHDHAEIQTAQRDLLEINFQARDEWFR 454
Qy 499 LQRKMNVDVSS 509
Db 455 VKDKNFDMSN 465

RESULT 9
Q9H0H6
ID Q9H0H6 PRELIMINARY; PRT; 823 AA.
AC Q9H0H6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Hypothetical protein DKFZp434A171.
GN Name=DKFZp434A171;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TiSSUE=Testis;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SSP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136793; CAB66727.1; -.
DR HSP; Q60778; IOY3.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS00088; ANK_REPEAT; 4.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 823 AA; 93968 MW; 9D2DD5F52C379B2 CRC64;

Query Match 20.0%; Score 519; DB 2; Length 823;
Best Local Similarity 30.1%; Pred. No. 2.1e-13;
Matches 148; Conservative 85; Mismatches 152; Indels 106; Gaps 11;

Qy 36 QKSVPNKALELKNEQT-----LRADILPSESQKQDYESSWDSLSLCT 80
Db 249 QQILEHKKILKKEKSDVGSSESASVSIFFHRLRVDSLPAASDDKOLNVATKQCVPEKSEP 308
Qy 81 VSQKDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKAE 140
Db 309 -----LPGSSHEKGNRIVNGQE-----G 327
Qy 141 PPEKPSAFPEAEMQKSVPNKALELKNEQTLRADQMPP--SESKQKVEENSWDSLSURE 198
Db 328 PPAKHPSLKPSSTEVEDPAVKGVQKQKQVQTLURAEQALPVASEEEOERHE-----RS 378
Qy 199 TVSQKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSCERARELOKDHCEORTGWE 258
Db 379 EKQPOVKEGNTNKSEKIQLSENICDSTS-----SAAAGRLT 416
Qy 259 QMKKFCVLKLLSEAKEIKSOLENOKVKEQELCSVRLTLNQBEKKRNADILNEKIRE 318

```

```

Db 417 QQRK-----IGKTYPOQFP-KKLKEBHDC-----TLKQNEEKTNNVLMYKKNRE 461
Qy 319 ELGRIBEQHRELVKQLEQALRIQDIELKSVESNLNOVSHTHENNYLHENCMLKE 378
Db 462 ELERKEQYKVEAK-QLEPTVOSLEMSKSTARNTPNRDPHNHEEMKGLMDENCILKAD 520
Qy 379 IAMLKLEIATLKHQVQOEKRYFEDIKILKEKNAELQMTLKEESLTKRASQYSGOLKV 438
Db 521 IAILRQELCTWMDNLEKENYKDKIVKTNNALEKYIKLNEEMITEAFRYQOELND 580
Qy 439 LIAENTMLTSK-LKEKQDKLEIAEIESHHPRLASAVODHQIIVTSRKSPAFHAGDA 497
Db 591 LKAENTRLNAELLKESKRLADIESYQSLAAATSKHSESVKTERNLKALERTRDV 640
Qy 498 CLQRKMNVDVS 508
Db 641 SVQVEMSSAIS 651

RESULT 10
Q81ZM7 PRELIMINARY; PRT; 453 AA.
AC Q81ZM7;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Melanoma-associated antigen (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bruno R., d'Orlando O., Altomonte A., Lamaj E., Maio M., Pucillo C.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF543495; AAN40505.1; -.
FT NON TER
SQ SEQUENCE 453 AA; 53140 MW; 81DF80012122BBCA CRC64;

Query Match 18.3%; Score 475; DB 2; Length 453;
Best Local Similarity 45.3%; Pred. No. 6.6e-12;
Matches 107; Conservative 54; Mismatches 62; Indels 6; Gaps 4;

Qy 275 KEIKSQLENOQKVKWQELCSVRLTLNQBEKKRNADILNEKIRELGRIEHQHRELVK 334
Db 2 BEIKSQLEHLELEKELCSLRFPAIQEKKRNVEEVHQVKREKLRTESQYRIEADVT 61
Qy 335 QQLEQALRIQDIELKSVESNLNOVSHTHENNYLHENCMLKKEIATLKHQYQ 394
Db 62 KPIKPAKSAEVELKTGGNSNQVSETDEKED-LLHENRLMQDEIARLREKDTIKQNQL 120
Qy 395 EKENKYFEDIKILKEKNAELQMTLKEESLTKRASQYSGOLKVLIAENTMLTSKL-KEK 453
Db 121 EK--KYLKDFEIVKVRKHEDLQALKRNGETLAKTIACYSGQAAUTDENTTLRSKLEQR 178
Qy 454 QKLEIAEIESHHPRLASAVODHQIIVTSRKSPAFHAGDAC--LQRKMNVDV 507
Db 179 ESRQRLETQMOSYHCLNARNACDHQSHSKEDQLAQGTVDYKCRHLQENLSHV 234

RESULT 11
O60311 PRELIMINARY; PRT; 641 AA.
AC O60311;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE KIAA0565 protein (Fragment).
GN Name=KIAA0565;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98230545; PubMed=9628581;
RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N.,
RA Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:31-39(1998).
DR EMBL; AB011137; BAA25491.2; -.
FT NON TER
SQ SEQUENCE 641 AA; 75777 MW; F11566F2E6A0B9F4 CRC64;

Query Match 16.6%; Score 431.5; DB 2; Length 641;
Best Local Similarity 40.2%; Pred. No. 5.2e-10;
Matches 106; Conservative 43; Mismatches 58; Indels 57; Gaps 4;

Qy 296 RLTLNQBEKKRNADILNEKIRELGRIEHQHRELVKQLEQALRIQDIELKSVESNL 355
Db 6 RLALKQNEEKKRNADMLYNKDSQRLRIKEBECGVETKQQLKWNLRRLVKELRTVRNL 65
Qy 356 -----NOVSHTHENNYLHEN 372
Db 66 DLVVQERNDQAQQLSEQDARILODQILTSKQLEMARKKMNSEISHRHQKEXDLPHED 125
Qy 373 CMLKKEIATLKHQVQOEKRYFEDIKILKEKNAELQMTLKEESLTKRASQY 432
Db 126 CMLQEEIATLKHQVQOEKRYFEDIKILKEKNAELQMTLKEESLTKRASQY 185
Qy 433 SGOLKVLIAENTMLTSKLKE-KQDKLEIAEIESHHPRLASAVODHQIIVTSR----- 484
Db 186 SGQLNLTAEINKILNSELNGKQERLEIMESYRCRLAAAVRDCQSQOTARDLKLDQ 245
Qy 485 KSQEPAFHAGDACLQRKMNVDVS 508
Db 246 RTQEWVR-----LHDKMKVDMS 263

RESULT 12
Q7K5Q6 PRELIMINARY; PRT; 2006 AA.
AC Q7K5Q6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 3.
GN Name=maebi; falciparum.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072559; PubMed=12076768; DOI=10.1016/S0166-6851(02)00067-1;
RA Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
RL Mol. Biochem. Parasitol. 122:35-44(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in
RT adhesive molecules of malaria parasites.";
RL Mol. Biol. Evol. 19:1128-1142(2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY042084; AAO73469.1; -.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
DR SEQUENCE 2006 AA; 237620 MW; EE75EF42E2BF767D CRC64;
SQ SEQUENCE 2006 AA; 237620 MW; EE75EF42E2BF767D CRC64;

```

Query Match 12.0%; Score 313; DB 2; Length 2006;
 Best Local Similarity 27.3%; Pred. No. 9.4e-05;
 Matches 138; Conservative 94; Mismatches 186; Indels 88; Gaps 22;

QY 19 KAEPEKPSAFPEAIEMQKSVNKALELN--EOTLRADAILPSSKQKDYESSWDSSES 76
 Db 1293 KAEKKKKADELKKAEEKK---KADELKKAEEKKKADEVKKAEEKKKADELKKAEEKK 1348

QY 77 LCETVSQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQOT 136
 Db 1349 KADE-----LKKSEKKKADELKKAEEKKKADELK---KAEEKKKADELK--- 1393

QY 137 FKAPEPEKPSAFPEAIEMQKSVNKALELN--EOTLRADQMPFSESQKQKVEENSWDSE 194
 Db 1394 -KAEKKKKADELKKAEEKK---KADELKKAEEKKKADEL---KKAEEKK-KAE 1441

QY 195 SLRETVSQK--DVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELOKHCEQ 252
 Db 1442 NLKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1501

QY 253 RTGKME--QMKKKFCVLKK-----KLSEAKEIKSOLENOKV--KWEQELCSVRL-- 297
 Db 1502 KADELKKAEKKKKADELKKAEEKKKADELKKAELKKAEEKKKADEVKQKREBERNNALRR 1561

QY 298 --TLNQ-----EEKRNADILNEKIREELGRIBEOHRLKLEVKQQLLEQALR 342
 Db 1562 AEILKQIEKKRIEVMKLYEEEKKMAEQLKKEEEKIK--AEQLKKEEEKKKVQQLKK 1619

QY 343 IQDIELKSVESNLQVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQVQEKENKYFE 402
 Db 1620 KEEEEKKKAQOLKKE-----EENKIKAEQLKKEEEKKKAEL--ELKKE-EEEEKKAE 1671

QY 403 DIKILKKEAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKELILAE 462
 Db 1672 QLKKEEEKKKVEQ-LKKKEEEKKKAEL---QLKKEEENKIKVQELKKEEEKKKAEL 1726

QY 463 IESHHPRLASAVQHDQIVTSRKSOE 488
 Db 1727 -----ELKKEEEKKKVQQLKKEE 1746

RESULT 13

QY7K5Q5 PRELIMINARY; PRT; 2019 AA.
 AC Q7K5Q5;
 DT 05-JUL-2004 (TremBLrel. 27, Created)
 DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TremBLrel. 27, Last annotation update)
 DE Erythrocyte binding protein 2.
 GN Name=maeb1;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22072559; PubMed=12076768; DOI=10.1016/S0166-6851(02)00067-1;
 RA Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
 RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
 RL Mol. Biochem. Parasitol. 122:35-44 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22077637; PubMed=12082132;
 RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
 RT "Evolutionary relationships of conserved cysteine-rich motifs in
 adhesive molecules of malaria parasites.";
 RL Mol. Biol. Evol. 19:1128-1142 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY042084; AAQ73468.1; -
 DR InterPro; IPR008602; Duffy_binding.
 DR Pfam; PF05424; Duffy_binding; 1.

SQ SEQUENCE 2019 AA; 239410 MW; A626F2C684C08785 CRC64;

Query Match 12.0%; Score 313; DB 2; Length 2019;
 Best Local Similarity 27.3%; Pred. No. 9.5e-05;
 Matches 138; Conservative 94; Mismatches 186; Indels 88; Gaps 22;

QY 19 KAEPEKPSAFPEAIEMQKSVNKALELN--EOTLRADAILPSSKQKDYESSWDSSES 76
 Db 1293 KAEKKKKADELKKAEEKK---KADELKKAEEKKKADEVKKAEEKKKADELKKAEEKK 1348

QY 77 LCETVSQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQOT 136
 Db 1349 KADE-----LKKSEKKKADELKKAEEKKKADELK---KAEEKKKADELK--- 1393

QY 137 FKAPEPEKPSAFPEAIEMQKSVNKALELN--EOTLRADQMPFSESQKQKVEENSWDSE 194
 Db 1394 -KAEKKKKADELKKAEEKK---KADELKKAEEKKKADEL---KKAEEKK-KAE 1441

QY 195 SLRETVSQK--DVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELOKHCEQ 252
 Db 1442 NLKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKK 1501

QY 253 RTGKME--QMKKKFCVLKK-----KLSEAKEIKSOLENOKV--KWEQELCSVRL-- 297
 Db 1502 KADELKKAEKKKKADELKKAEEKKKADELKKAELKKAEEKKKADEVKQKREBERNNALRR 1561

QY 298 --TLNQ-----EEKRNADILNEKIREELGRIBEOHRLKLEVKQQLLEQALR 342
 Db 1562 AEILKQIEKKRIEVMKLYEEEKKMAEQLKKEEEKIK--AEQLKKEEEKKKVQQLKK 1619

QY 343 IQDIELKSVESNLQVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQVQEKENKYFE 402
 Db 1620 KEEEEKKKAQOLKKE-----EENKIKAEQLKKEEEKKKAEL--ELKKE-EEEEKKAE 1671

QY 403 DIKILKKEAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKELILAE 462
 Db 1672 QLKKEEEKKKVEQ-LKKKEEEKKKAEL---QLKKEEENKIKVQELKKEEEKKKAEL 1726

QY 463 IESHHPRLASAVQHDQIVTSRKSOE 488
 Db 1727 -----ELKKEEEKKKVQQLKKEE 1746

RESULT 14

QY8IHP3 PRELIMINARY; PRT; 2055 AA.
 AC Q8IHP3;
 DT 01-MAR-2003 (TremBLrel. 23, Created)
 DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE MAEBL, putative.
 GN ORFNames=PF11_0486;
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2255705; PubMed=12368864; DOI=10.1038/nature01097;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.S., Nene V., Shallow S.J., Sub B., Peterson J., Angiuoli S.,
 RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 falciparum.";
 RL Nature 419:498-511 (2002).
 DR EMBL; AE014843; AAN36066.1; -
 DR HSSP; P04268; 1IC2.
 DR GO; GO:0016021; C:integral to membrane; IEA.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:56:15 ; Search time 14.4804 Seconds
(without alignments)
3402.035 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDMTFKA.....IAGDACLQRXNDVDSSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	265	10.2	2954	T14156	kinesin-related pr
2	264	10.2	1738	T14867	interaptin - slime
3	261.5	10.1	2116	A26655	myosin heavy chain
4	255	9.8	2017	A36014	myosin heavy chain
5	255	9.8	2057	S61477	myosin II heavy ch
6	250	9.6	944	S26710	spindle pole body
7	247.5	9.5	2253	T30336	nuclear/mitotic ap
8	244.5	9.4	1979	C71622	hypothetical prote
9	241.5	9.3	1313	F96673	hypothetical prote
10	240.5	9.3	946	S28061	SCP1 protein - rat
11	239.5	9.2	1827	T16270	hypothetical prote
12	239	9.2	1790	S67593	transport protein
13	239	9.2	3259	A56539	giantin - human
14	238	9.2	992	T46337	hypothetical prote
15	238	9.2	3225	I52300	giantin - human
16	237.5	9.1	1993	S49461	synaptosomal compl
17	236	9.1	1690	T13030	microtubule bindin
18	235.5	9.1	1642	T08880	NMDA receptor-bind
19	235.5	9.1	2245	T18278	myosin heavy chain
20	232	8.9	544	I36911	involucrin L - dou
21	232	8.9	1388	T08633	serine/threonine-s
22	230.5	8.9	1432	B85431	trichohyalin like
23	230.5	8.9	1939	A46762	myosin alpha heavy
24	230	8.8	2442	T08621	centrosome associa
25	230	8.8	2663	S28261	centromere protein
26	229	8.8	1164	T24806	hypothetical prote
27	229	8.8	1875	T38173	myosin-like protei
28	229	8.8	1979	S03166	myosin heavy chain
29	228.5	8.8	978	A70387	conserved hypothet

ALIGNMENTS

RESULT 1

T14156

kinesin-related protein - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T14156

R:Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.

Cell 91, 357-366, 1997

A:Title: CENP-E is a plus end-directed kinetochore motor required for metaphase chromosome

A:Reference number: 217893; MUID:98028574; PMID:9363944

A:Accession: T14156

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2954 <WOO>

A:Cross-references: UNIPROT:O42263; EMBL:AF027728; NID:G2586070; PID:G2586071; PIDN:AAAC

C:Genetics:

A:Gene: XCENP-E

C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match 10.2%; Score 265; DB 2; Length 2954;
Best Local Similarity 21.8%; Pred. No. 0.00033;
Matches 106; Conservative 108; Mismatches 172; Indels 100; Gaps 15;

Qy	5	IPTKALELMDMTFKAPPE--KPSAFEPAPKQSVNPKALELKNQOTLRADELTPSES	62
Db	1888	IETLSLSLKEKE-FALQEAQKADAAKTIIDITEKISNIEEQLLQQTNLKELYERES	1946
Qy	63	KQDYESSWDSLSCTVSQKQVCLPKATHQKE-----IDKINGKLESPDNDGF	113
Db	1947	LIQCKEQALNTEHLRETLKSKDLALGMEQERDEAANKVIALTEKMSLSEQINENVT	2006
Qy	114	LKAPCRMKVSIPTKALELMDMTFKAPPEKPSAFEPAPKQSVNPKALELKNQOTLRA	173
Db	2007	LKEGEGEK-----ETFLQRPSSQKQSSQMEELRESLTKDLQ-----	2045
Qy	174	QMPFPSSKQKVBENSWSDE-----SLRETSQKQVCPKATHQKEMDKISGKLEDS	226
Db	2046	-----FEAEKEISEATNEIKNLTAKISSLEEILQNASILNEAVSERENLRHS-QQILV	2098
Qy	227	TSLSKILDVTHSCBRARELQKHCEQRTGRKMQMKKFCVLKXKLSEAKEIKSOLENQKV	286
Db	2099	SELQSLSTLKSRDHFAQSKREKDEAVNKIASLAEEIKILTKEWDFRDSKESLQEQSS	2158
Qy	287	KWQELCSVRLTNQEEKRNADILNEKIFREELGRTE--OHRKELEVQKQLEQALRIQ	344
Db	2159	HLSELCTYKTELQMLKQKQED---INNKLAEKVEDELLQH-----	2198
Qy	345	DIELKSVESNLNQVSHNENYLLHENC-----MLKKEIAMLKLEIATLKHOYEKENKY	400
Db	2199	---LSSKEQJQIQMELRNEKLNVELCEKXMDMEKISVRL-----MQNEPQSEEDV	2251
Qy	401	FEDIKILKEKNAELQMTLKLKESLTKRASQYSGQLKVLIAENTMLTSLKJKEKQDKEILE	460

Db 2252 AERMDILSRNOEIQ-----ELMEKISAVYSEQ-----HTLLSLSSSE-----LQ 2291
QY 461 AEIESH 466
Db 2292 KETEAH 2297
RESULT 2
T14867
interactin - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14867
R:Rivero, F.J.; Kuspa, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.
J. Cell Biol. 142, 735-750, 1998
A:Title: Interactin, an actin-binding protein of the alpha-actinin superfamily in Dictyo
te.
A:Reference number: Z18248; MUID:98365468; PMID:9700162
A:Accession: T14867
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1738 <RIV>
A:CROSS-references: UNIPROT:O76329; EMBL:AF057019; NID:G3549260; PID:G3549261; PIDN:AAC3
C:Genetics:
A:Gene: abpD
A:Introns: 173/2; 1680/1
Query Match 10.2%; Score 264; DB 2; Length 1738;
Best Local Similarity 22.1%; Pred. No. 0.0002;
Matches 133; Conservative 120; Mismatches 203; Indels 146; Gaps 25;
QY 10 LELMDQTPKABPPKPSAFEP-----AIEMOKSVPNKALEKNEQTLRADE 56
Db 585 LQLQDQDIK----EKEFQFERQQLLSQIDSITTNIOEQYDFRNNLQOEFNTQOOLNQE 640
QY 57 I--LPSESKQ--KDYEESSWD-----SESLCETVSQKDVCLPKATHQKE---ID 98
Db 641 THRLTQQLQIINTDYNEKQTQOSEIKQNTQINEQLNKSQKDEIEKLSNQOQQODE 700
QY 99 KINGLEESPDNDGFLKPCRMKVSIPTRKALE-----LMDMTFFKABPEKPSA 147
Db 701 KINLLLEIKEDCLIE-----RINQQLLENIDLSKYQQLLEFENFK-----744
QY 148 FEPALIMOKSVPNKALEKNEQTLRADQMP-----SESKQKVEENSW 191
Db 745 ----LNSKSEKENQLNELQSQDERFNLNDEKLEKEKQLOQIEDEFNQYKQOQSSNEN 800
QY 192 DSLSRETYSQKDVCPKATHQKEMD--KISGLEDSTSLKILDTVHSCERARELQKH 249
Db 801 IDQOLQSTIIE----LSELKEQELNDSKLEKEKQLOLQOEFQDLN-----EKNQKH 851
QY 250 CEQRTGMEQMKKFCVLKKLSEAKEIKSLENQKVK-----EQLCSVRLT 298
Db 852 QDQ----LELLEKQLOQOEDYQDLNETNQSIENQNLNKENLNEKEQELLKLQ 907
QY 299 LQNEEEKRR-----NADIINEK-----IREELGRIEEQHRKELEVKQLEQA 340
Db 908 LNOQIEKIQDQOEFQKQNSINTELNVNKNKELIQDYDQLKQNNR--DEKEND 964
QY 341 LRQDTELKSVESNLNQVSHTHEN---ENYLLH---ENCMLKEIAMLKLEIATLKHQYQ 394
Db 965 LIEKENQLKSIQNELNQLNLIKESDHKEQOLKQOSIENDLIEKENIQLOLQ-SOLNEQ 1023
QY 395 EKENKYFE---DITKILKERNAEIOMTLKLEESLTKRASQYSQOLKILTA-----ENTMLT 447
Db 1024 QQSOLSEKQDQOLNLIKQNFQDQKQOQKQOSIENDLIEKENIQLOLQOLNEQ 1083
QY 448 SKLKEK--QDKETILEA--EIESHPRLASAVQDHDQIVTSRKSQEPAPHIAGDACLRKWN 504
Db 1084 NQUSEKQDQOLNLIKQNESQKQOLKQOSIENDLIEKENIQLOLQOLNEQQLQSEVS 1143
QY 505 VD 506

Db 1144 ID 1145
RESULT 3
A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Dictyostelium discoideum
C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C:Accession: A26655; A24728; S00250
R:Warrick, H.M.; De Lozanne, A.; Scheel, J.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium disco
A:Reference number: A26655; MUID:87092266; PMID:3540939
A:Accession: A26655
A:Molecule type: DNA
A:Residues: 1-2116 <WAR>
A:CROSS-references: UNIPROT:P08799; GB:M14628; GB:M1938; NID:G167834; PIDN:AAA3227.1;
R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Weinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A:Reference number: A24728; MUID:86016788; PMID:3901008
A:Accession: A24728
A:Molecule type: mRNA
A:Residues: 2035-2116
R:Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A:Reference number: S00250; MUID:88112226; PMID:2828113
A:Accession: S00250
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1734-1893 <WAG>
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phos
F:1-818/Domain: globular head <HEAD>
F:179-186/Region: myosin motor domain homology <MMOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:819-2116/Domain: alpha-helical rod <ROD>
Query Match 10.1%; Score 261.5; DB 2; Length 2116;
Best Local Similarity 22.0%; Pred. No. 0.00032;
Matches 135; Conservative 107; Mismatches 212; Indels 161; Gaps 24;
QY 8 KALELMDQTPKABPPKPSAFEPALIMOKSVPNKAL-----ELKNEQTLRADEILPSESK 63
Db 894 RELRIR-VEDMESLDEKLALENLQKRSVEEKVRDLEELQEQKLRN---TLEKL 948
QY 64 QKDYEESSWD-----SESLCETVSQKDVCLPKATHQKEIDKINGKL-BESPNDGFLKAPC 118
Db 949 KKQYEEELKEEMKRVNDGQSDTISRLEKI--KDELQKEVEELTESFSESKDKGVLEKTRV 1006
QY 119 RMKVSIPFKALELMDQTFKAE-----PPEKPSAFEPALIMOKSVPNKAL-- 163
Db 1007 RLQSELDLTVRLDSETKDKSELLRQKKLEEBELKQVQELAAETAAKLAQEAANKLQ 1066
QY 164 -----ELKNEQTLR----- 172
Db 1067 EYTELNEKFNSEVTARSNVESKKTLESQLVAVNNELDEENKRDALKKKKALDAMLEE 1126
QY 173 -ADQMFPSESKOK-----KVSENSWDSLSRETYSQKDVCP-----KATHQEMDKISG 221
Db 1127 MKDQLESTGGKSKSYDLKVKQES--DMEALRNQISLSELOSTIAKLEIKSTLSEGEVARIQG 1185
QY 222 KLEDSTSLKILDTVHSCERARELQKHCEORTGME--OMKKKFCVLKKLSEAKEIKSQ 280
Db 1186 ELE-----AEQLAKSNVEKQKKVDELDEDK-----SAQLAEETAQA 1224
QY 281 LENQKVKWEQELCSVRLTNQEEFKRRNADILNEKIREBELGRIBEQHKELEVKQOLEQA 340
Db 1225 LDKLKKLEQELSEVOTQLSEANNKVNNSDSTNKHLETSPNNL-----KLELEAEQAKQA 1280

Qy 341 LRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLKUEIATLKHQOE----- 395
Db 1281 LEKKRLGL---ESELKHVNEOLEBEKKQKESN---EKRKVDLEKEVSELKQIEEVASK 1334
Qy 396 -----KEN-----KYFEDIKILKEKNAELOMTLKLKEESTLTKRASQYSQOLKVL 440
Db 1335 KAVTEAKNKKSESELDEIKRQYADVSSKDSKVEQLTKLQAKNEELRNNTAEAEAGOLD--- 1391
Qy 441 AENTMLTSLKLEKODKEILEA-----ETESHHPRLASAVQDHQIVTSRSKQ-SPAPHIA 494
Db 1392 ----RAERSKKAEFDLEAEAVKNELEETAKKVAEKAMKKAETDYRSTKSELDDAKNVS 1446
Qy 495 GDACLO-RKNNDVVS 508
Db 1447 SEQYVQIKRLNEELS 1461

RESULT 4
A36014
myosin heavy chain, nonmuscle - fruit fly (Drosophila melanogaster)
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Drosophila melanogaster
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A36014; B36014
R:Ketchum, A.S.; Stewart, C.T.; Stewart, M.; Kiehart, D.P.
Proc. Natl. Acad. Sci. U.S.A. 87, 6316-6320, 1990
A:Title: Complete sequence of the Drosophila nonmuscle myosin heavy-chain transcript: cDNA
A:Reference number: A36014; MUID:90349606; PMID:2117279
A:Accession: A36014
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2017 <RET>
A:Cross-references: UNIPROT:Q99323; GB:M35012
C:Genetics:
A:Gene: FlyBase:zip
A:Cross-references: FlyBase:FBgn0005634
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated
F:1-2017/Product: myosin heavy chain, form I #status predicted <MA1>
F:46-2017/Product: myosin heavy chain, form II #status predicted <MA2>
F:135-815/Domain: myosin motor domain homology <MMOT>
F:225-232/Region: nucleotide-binding motif A (P-loop)
F:598-631/Region: actin binding #status predicted
F:705-727/Region: actin binding #status predicted
F:888-2017/Domain: coiled coil #status predicted <COI>
F:888-1328/Region: S2
F:1329-2017/Region: light meromyosin
F:176/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F:231/Binding site: ATP (Lys) #status predicted
F:745,755/Active site: Cys #status predicted

Query Match 9.8%; Score 255; DB 1; Length 2017;
Best Local Similarity 22.2%; Pred. No. 0.00059;
Matches 122; Conservative 113; Mismatches 182; Indels 132; Gaps 23;

Qy 14 DMOTFKAPPEKPSAFEPAPTEMOKSVPNKALELNE-----QTLRADRLPESKQ----- 64
Db 1170 DLEAKA-----ARAKAEKVRDLSEELAKNELLDLSLDTTAAQQLSKRQELAT 1222
Qy 65 --KDYESSWDSSELCTVSKQVCLPKATHQEKIDKINGKLE----- 105
Db 1223 LKKSLEETVNHGVLADMRHK-----HSQELNSINDQLENLRKAKTVLEKAKGTLE 1274
Qy 106 -ESPD-----NDGFLKAPCRMVSPPTKALELMDMTQFKAEPPKPSAFE- 149
Db 1275 AENADLATELSVNSRQENRRRK---QAESQIAELQVKLAETARASELQEKTKLQ 1331
Qy 150 -----PAIEMOKSVPNKALELKNQTLRADQMPFSESKQK-----KVEENSWDSES 195
Db 1332 EAENITNLEAEALKASAAVKSANMESQLTEAQQLLEETFRQKLGSLSKLRQIESEKEA 1391
Qy 196 LRETYSQKDVCPKATHQK-----EMDKTSGKLESDTSLGSKLDTVHSCERARELQ 248

Db 1392 LQEQLEEDDEA--KRNRYERKLAEVTTQMQEIKKAAEEDADLAKELE-----EGKKRLNKD 1444
Qy 249 HCEQRTGMEOMKKKFCVLKKKLLSEA-KEIKSQLENOKVQWQELCSVRLTLNQOEEKRR 307
Db 1445 -----IEALERQVVELTAQNDRLDKSKKIOSELEDATIELEAQ-----RTKVELEKKQK 1495
Qy 308 NAD-----ILNPKIREELGRIE-----EQRKSELEVQKQLEQAL-RTODIE--LKSVES 353
Db 1496 NFDKILAEKAKISQIAQERDTAREAREKETKVLVSRELDEAFDKIEDLENKRKTLQN 1555
Qy 354 NLNQVSHTHENENYLLHNCMLKKEIAMLKUEIATLKHQOEKEN--KYFEDIKILKEKN 411
Db 1556 ELDDLANTQGTADKNVHE---LEKAKRALESQALKAQNEELEDLQUTSDAKLRLVN 1612
Qy 412 ----AELQMTLKLKESLTKRASQYSQOLKVLIAE-----NTMLTSLKLEKQDKKEI 458
Db 1613 MQALRSQFERDLLAKEGAEKRRGLVKQLRDLTELDEERKQRTAAVASKKLEGGDLKE 1672
Qy 459 LEABIESHH 467
Db 1673 IETTMEHNN 1681

RESULT 5
S61477
myosin II heavy chain, non-muscle - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: S61477; S65349
R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.
J. Mol. Biol. 255, 98-109, 1996
A:Title: Molecular organization and alternative splicing in zipper, the gene that encodes
A:Reference number: S61477; MUID:96144835; PMID:8568878
A:Accession: S61477
A:Molecule type: DNA
A:Residues: 1-2057 <MAN>
A:Cross-references: UNIPROT:Q94987; EMBL:U35816
R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.
submitted to the EMBL Data Library, September 1995
A:Reference number: S65349
A:Accession: S65349
A:Molecule type: DNA
A:Residues: 1-1908, 'NL', 1911-2057 <MAW>
A:Cross-references: EMBL:U35816; NID:g1141789; PIDN:AA809049.1; PID:g1572481
C:Genetics:
A:Gene: zip
A:Cross-references: FlyBase:FBgn0005634
A:Introns: 10/3; 90/3; 162/3; 215/1; 255/3; 263/3; 303/3; 334/3; 1248/1; 1349/3; 1526/3;
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: alternative splicing; ATP; nucleotide binding; P-loop
F:135-855/Domain: myosin motor domain homology <MMOT>
F:225-232/Region: nucleotide-binding motif A (P-loop)

Query Match 9.8%; Score 255; DB 2; Length 2057;
Best Local Similarity 22.2%; Pred. No. 0.00061;
Matches 122; Conservative 113; Mismatches 182; Indels 132; Gaps 23;

Qy 14 DMOTFKAPPEKPSAFEPAPTEMOKSVPNKALELNE-----QTLRADRLPESKQ----- 64
Db 1210 DLEAKA-----ARAKAEKVRDLSEELAKNELLDLSLDTTAAQQLSKRQELAT 1262
Qy 65 --KDYESSWDSSELCTVSKQVCLPKATHQEKIDKINGKLE----- 105
Db 1263 LKKSLEETVNHGVLADMRHK-----HSQELNSINDQLENLRKAKTVLEKAKGTLE 1314
Qy 106 -ESPD-----NDGFLKAPCRMVSPPTKALELMDMTQFKAEPPKPSAFE- 149
Db 1315 AENADLATELSVNSRQENRRRK---QAESQIAELQVKLAETARASELQEKTKLQ 1371
Qy 150 -----PAIEMOKSVPNKALELKNQTLRADQMPFSESKQK-----KVEENSWDSES 195
Db 1372 EAENITNLEAEALKASAAVKSANMESQLTEAQQLLEETFRQKLGSLSKLRQIESEKEA 1431

Matches 123; Conservative 105; Mismatches 178; Indels 126; Gaps 21

QY 37 KSVNKALELKNQOTLRADIELPSSKQKDYEESSWDSSESLCETVSKQDVCL--PKATHQ 94
 Db 244 KTVKDQVLENNSDVQ---LKRSEKDELKNLWNLNELKSAEAKDTQLEFKKNELR 300
 QY 95 KEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKAL--ELMDMQTKABPPKPSAFPAI 152
 Db 301 KRTWELNELKIKSDMDQLK----QKQNESKRLKDELNELETKFSNGSQSSAKENEL 355
 QY 153 EMQKSVPNKALELKNQOTLRADQMFPSSEKQKQVENSWSSESLRETYSQKDVCPKATH 212
 Db 356 KMLK--NKIAELEEEETSTKNSQLIAKEGKLASIMAUQTQLES---KLNQKDSQJ--GSR 407
 QY 213 QKEMDKISGLESTDSKILDTVWSCERARELQKHCHCEQRTGKMEQMKKFCVLKKKLS 272
 Db 408 EEEUKYNDKLQKDIRAR--EETVSKDERIIDLQK-----KVKQLENDLFIKKTHS 458
 QY 273 EAK-----ETKSQ-----LENQ-KV-----KWEQLCSVRLTNQBEKKERNADILNE 314
 Db 459 ESKTITONELESKDKLIKILENDLKVAQEKYSMEKEL-----KEREFNKYSIS 508
 QY 315 KIREELGRIEQHRKELEVQKQALRIQDIELKSVESNLNVQVSHHTENENYLLHNCM 374
 Db 509 KLDESKTYLTKISNLAAENSQKN--KIED-----NSTATHMKENY----- 549
 QY 375 LKKEITAMLKLEIATLRKHQYQENKNYPE-----DIKLKKNAEILQM 416
 Db 550 -EKQLESRLKDI EYKESAKDSBKTEELKIRIAENSASVSEKSKDIIKQKDEQISDUTQ 608
 QY 417 TLKKEESLTKRAS-----QYSGQLKVLIAENTMLTSKLKEKQDKE 457
 Db 609 NUKQDEISISLKIIDRYKDFNQLKSEGSNTQHDNLQILNLENKLI SEDELKLSLRD 668
 QY 458 ILEAIEIHHRLASAVQDHDQIVTSRKSQEPFHITAGD-----ACLQRKMN 504
 Db 669 SQKIEIENWKYNNLSLENDRLLTKEKES-----ASDKEREISILNRKLD 713

RESULT 7
 T30336
 nuclear/mitotic apparatus protein - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T30336
 R:Merdes, A.; Ramyar, K.; Vechio, J.D.; Cleveland, D.W.
 Cell 87, 447-458, 1996
 A:Title: A complex of NuMA and cytoplasmic dynein is essential for mitotic spin
 A:Reference number: Z20828; MUID:97053784; PMID:8898198
 A:Accession: T30336
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-2253 <MER>
 A:Cross-references: UNIPROT:P70012; EMBL:Y07624; NID:g1514670; PIDN:CAA68905.1
 C:Genetics:
 A:Gene: NuMA

Query Match 9.5%; Score 247.5; DB 2; Length 2253;
 Best Local Similarity 20.4%; Pred No. 0.0014;
 Matches 136; Conservative 118; Mismatches 235; Indels 177; Gaps 23

QY 7 TKALELMDMQTKABPPKPSAF--EPALEMQKSVPNKALE-----LKNEQTLRADEIL 58
 Db 1064 SRSLDILALKEGEVERLNKEALROEETIQOQQTITKLTEETALAKDKVALQEKIK 112
 QY 59 PSESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKID-----KINGKLEESPDNDG 112
 Db 1124 QQVQATGKAKEKAWALKSVISEKSRISCLSEQDTONKRDLSLCIQEQHQSGLGESQGLA 118
 QY 113 F--LKKAPCRMKVSI-----PTKALELMDMQTKABPPKPSAFPAI----- 153
 Db 1184 LIADLEKKCKEOKELICEAQNKAAAKTILASEKASVSRQLGQTQALEIFIGKERQACD 124

Qy 154 -----MKSVPNKALEKNEOTLRADQ-----MPPSESKOKK 185
Db 1244 LOKOLESWA VOEEKETELQALQKELFKHVKQOELOSQTSFTDSGEALLYLSEAQERQQA 1303
Qy 186 VEENSWSDESLETSQKDVCPKATHQKEMDKISGLKEDSTSL-----KILDTVHSC 239
Db 1304 LTKAEQAEQYQKEIEMKNKVENSL--QAEIKILSSKVTYNEEVSVDPEORLLKETSXA 1361
Qy 240 ERARELOKDHCEQRTGRMEOMKKPCV-----LKKKLSEAKE 276
Db 1362 KLEEKMQKLHMELEASFKELLEKNCAIDCLTTEAQNKLGEADQORMAVDSLQOKLSSKAE 1421
Qy 277 IKSQLEKNQKVM-----EQELCSVRLTL--NQE-----BEKERNADILNEK 315
Db 1422 TNHTLQOEIQAWQKNCAEKEOQICSLOQNLKSNOSLLEEFASLKHYSQEIITAEADLMQEK 1481
Qy 316 IREEL---GRIEEOHRKELEVKQO---LEQALRIQDIELKSVE-----NLNOVSHT 361
Db 1482 HQBELLSHKKLTERRQAELEKAKEDMTIVILLKEKLNQELQHKFQSENSYSLTQISHL 1541
Qy 362 HENENYLLHENCMLK-----KEIAMKLBIATLKHQYOB-----KENKYPED 403
Db 1542 QOVNSQLLQANQSLQISDOGAKKLESEMSTLKEQHKEEMKTLRLQYBKTTLREGNKQVOE 1601
Qy 404 I-----KILKEKNA--BLQWTLKLEESLTKRASOYSGQLKVLIAENTM 445
Db 1602 TSLOLETVTSKYDVHKSVLQDKQTFQBEKQRLQLQVQELNKLQSL-----QEKTI 1652
Qy 446 LTSKLKEKQDKEILLEAEIESHHPR---LASAVODHDQIVTSRKSQEPAPFIAGDACLQK 502
Db 1653 RSQOQKLQREGTHEEADKSHKRVLELESQLEQQTQAVEHYKAKQKAVHYDA--KKK 1710
Qy 503 MNVDVS 508
Db 1711 QNQEELS 1716

RESULT 8
C71622
hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 09-Jul-2004
C:Accession: C71622
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: C71622
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1979 <GAR>
A:Cross-references: UNIPROT:O96133; GB:AE001375; GB:AE001362; NID:g3845105; PIDN:AAC7181
A:Experimental source: clone 3D7
C:Genetics:
A:Gene: PFB0145c

Query Match 9.4%; Score 244.5; DB 2; Length 1979;
Best Local Similarity 23.2%; Pred. No. 0.0017;
Matches 141; Conservative 92; Mismatches 183; Indels 191; Gaps 30;
Qy 34 EMQKSVNKALEKNE--QTLR-----ADEILPSEK-----QKDYEE 69
Db 118 EISKNIHKNELENQKLDTLKSISSLNKIVNYESKIELEKELVKDKNDNDN 177
Qy 70 SSWSESCEVTSQK-DVCLPKAT--HOKED-----KINGK----- 103
Db 178 KLEKE---DFVQKIDMLNEKLLQKELDINKREKKNINEKKNIIKKETFTNIEKE 234
Qy 104 -LEESPNDGFLKAPCRMKVSIPTKALELMD---MOTFKAEPPKPSAREPATEMQKSV 158
Db 235 YLEKNKERE-----TISIEIIDIKHLEKLEIKIEKKEDLENLKNLLSK 280

Qy 159 PNKALELK-----NEQTLRA--DOMPPSESKOKKVE-----ENSWDSSES 195
Db 281 ENVLKELKGCVKCKNETINSNDNIIEKEKKYKLLVELEBKKNQIDLLNKOEKEKEK 340
Qy 196 LRETVSOKDVCVPKATHQKEMDKISGLKEDSTSLKILDTVHSCB--RARELOKQD-----H 249
Db 341 EREKEKEK-----KEKEKYDTLIKELKDEK--ISILEKVHSIKVREMDIEKREHNFJH 393
Qy 250 CEQRTGRMEOMKKPCF-----VLKKKLSEAKIKSQLENQKVKWEQELC-----SVRL 297
Db 394 MED---QLKDLKSNFVKNNQKLVYKCEIKNLKTELE---KKEKELADIENVSKEINK 446
Qy 298 TLNGBEKKRRNADILNEKIREE-----LGRIBEQHRKEL--EVQK-----QLEQ 339
Db 447 LINQNEKEKQILAFNKNHKEEIHGLKEELKESVKITKIETQELQEMVDIIOKELDQLOE 506
Qy 340 ALRIQ-----DIELKSVESNLNOVSHTHENE-----NYLL 369
Db 507 KYNQAIQISISIELEKKEKEYNQYKNTYIEEINNLEKLEETNKEYTNLQNYTNEINMLN 566
Qy 370 HENCMKKEIAMKLEIATLK---HOYOEKENKYFEDIKILKEKNAELQ---MTLKJKEE 423
Db 567 NDIHMLNGNIKTMTQISTLKNQVHLLNEQIDKLNNKGTLSKISLNLVQIMDLKBEKD 626
Qy 424 SLTKRASOYSGQLKVL-----AENTMLTSKLKQKQKEILEABIESHHPRLASAVQHD 478
Db 627 FLNNQIVDLSNQIDLLTRKBEKENKMLEQENKYQEMELLRGNIKS-----SENILNND 681
Qy 479 QIVTSRK 485
Db 682 BEVCDLK 688

RESULT 9
F96673
hypothetical protein F13011.30 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F96673
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizak, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1313 <STO>
A:Cross-references: UNIPROT:Q9XIP6; GB:AE005173; NID:g5042434; PIDN:AAD38273.1; GSPDB:G1
C:Genetics:
A:Gene: F13011.30
A:Map position: 1

Query Match 9.3%; Score 241.5; DB 2; Length 1313;
Best Local Similarity 24.5%; Pred. No. 0.0015;
Matches 146; Conservative 99; Mismatches 185; Indels 165; Gaps 28;
Qy 32 AIEMQKSVNKALEKNE---QTLRADE---ILPSEKQKDYESSWDSSELCTVTSQKD 85
Db 579 SMKLKESLLKBEELKNVTABISSIREWEGSVLEKIELSKVKESLVVDKFTKLQSIITGEA 638
Qy 86 VCLP--KATHQKED-----KINGKLEESPD-----NDGFLKAPCRMKVSIPT 125
Db 639 BELAGREAAHMKQIEELSTANASLVDEATKLQISVQESDELKEKEAGVYKIEELSVANE 698
Qy 126 TKALELMDMOTFKAP---PEKPSAFEPAIEMQKSVNKALEKNEQTLRADQWPP----- 178

Db 699 SLADNVTDLSQIVQESKDLKEREVAYLKIE-ELSVANESLVKETHKLOHIDQAEALRG 757
Qy 179 -SESKQKVEENSWDSRLSTVS-----QXDVCPKATHOKEMDKIS---GKLEDS 226
Db 758 REASHLKIEELSKENENLVNVMQNIABESKDLREREVAYLKIDELSTANGTLADN 817
Qy 227 TS-----LSKILDTVHSCEAR-----243
Db 818 VTNLQNISEENKELRERETITLKAAEELSELNVLVDKASKLOTVQVQENELRERETAYL 877
Qy 244 ----ELQKH---CEORT-----GKWEOMKKKCVLKKLSEAKEIKSOLENKVQWEQ 290
Db 878 KKEIELSKHILSDQETKLOISNHEELKERETAYLKIEELSKVQVEDLLNK-----EN 933
Qy 291 ELCVRLTLNOBEEKRRNADILNEKRIEELGRIBEOHREKELEVKQOOLEQALRIODIELKS 350
Db 934 ELHGAV-----FIEDLRKSDSLAOKKIEELSNF-----NASLLIKENELQAVVCNEBELKS 985
Qy 351 VE-SNLNQVSHTHENENYLH-----ENCMKKKEIAMIKL-----ETATLKHQVQE 395
Db 986 KQVSTLKTIDELSDLKQSLIHKKELOAAIVENEKIKAE-AALSLQRIEELTNLKTQTLID 1044
Qy 396 KENK---VFEDIKILKKNAELOMTLK-----LKEESLTKRASQVSGQLK---437
Db 1045 KQNELQGVFHENEELKAEAS---SLKKIDELHLLEQSWLEKSESEFQVTVQENLELKTQD 1101
Qy 438 VLIAENTMLTSKIEKQDKKEILEABIESHHPRPLASAVODHDQIVTVSRKSQEPAPH 492
Db 1102 ALAAKKIEELSKLKE---SLLEKETEL-KCREAAALE-----RWEEFSPKH 1142

RESULT 10
SCPI protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 08-Oct-1999
C:Accession: S28061
R:Meuwissen, R.L.J.; Offenberger, H.H.; Dietrich, A.J.J.; Riesewijk, A.; van Iersel, M.; HEMO J. 11, 5091-5100, 1992
A:Title: A coiled-coil related protein specific for synapsed regions of meiotic prophase
A:Reference number: S28061; MUID:93099884; PMID:1464329
A:Accession: S28061
A:Molecule type: mRNA
A:Residues: 1-946 <MEU>
A:Cross-references: EMBL:X67805; NID:G57212; PIDN:CAA48006.1; PID:G57213
C:Genetics:
A:Gene: SCPI
C:Keywords: DNA binding

Query Match 9.3%; Score 240.5; DB 2; Length 946;
Best Local Similarity 22.4%; Pred. No. 0.0011;
Matches 139; Conservative 103; Mismatches 193; Indels 185; Gaps 28;

Qy 13 MDMQTPKAEPPKPSAFPAIEAMQKSPVKNALKEQLTADRADEILPSSKQD-----YE 68
Db 187 LEHFKLKBDEHKEIQHLEE--EYQKEVNNK-----ENQVSLLIQSTKENKMKDLTFLE 240
Qy 69 ESSWDSSELCTVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKPCRMKVSIPTKA 128
Db 241 ESRDKANQLEBKTQDENLKELEKNEK--DHLTSELED-----IKSMQORSSTQK 289
Qy 129 LELMDMO---TFKAEPPKPSAFPAIEAMQKSPVKNALKEQLTADRADEILPSSKQD-----EQLTRADQM 176
Db 290 TLEEDLQIATKIYVLTETEEKAQNE---ELNKAHTHSLVTVTELKATCTTLEELLR----342
Qy 177 FPSSKQKQVENSWSDESIRETVSKQDVCPKATHOKEMDKISG-----KLEDSLSL---229
Db 343 ---TEQORLENNE-DQLKLTIMELQK-----KSELEEMTKFNKNKEVELESLKILAE 392
Qy 230 -SKILDTVHSCEAR-ARELOKHCE-----QRTGKMEQMKKFCVLKKK-----LSBAKEI 277
Db 393 DOKLLDEKKQVEKLAELQKGEQELTFLLQTRKEIHDLEVOVTVTKTSEHYLKQVEEM 452

Qy 278 KSOLENKVQKQOELCSVRLTLNOE-----EKRRNADILNEKIRBE 319
Db 453 KTELEKEKLK-----NIELTANSMLLENKCLVQEASDMVLELTKHQEDIIINCKQBE 506
Qy 320 --LGRIBEOHREKELEVKQOOLEALR--IQ-----DIELK 349
Db 507 RMLKQIETLEEKEMNLRLDELSVRKEFTIQGDEVCKLDKSEENARSIEYEVVLKKEKQMK 566
Qy 350 SVESNLN---QVSHTHENENYLLHENCMKK-----EIAMLKLEIATLKH 391
Db 567 ILENKNNLKKOIKENSKNIBELHOENKALKKSSAENKQNLNAYELKYNKLELASTQK 626
Qy 392 QYQEKENKFFEDIKILK-----EKNABLOMTLKLKEESLTKRASQVSGQLKVLI 440
Db 627 KKEEMINNYQKEIIEIKKISEEKLGEVEKAKATVDEAVKLOKE-IDLRQCHKIAEMVALM 685
Qy 441 AENTMLTSLKKEKQDKKEI-----LEAEIESHHPRLAS-----472
Db 686 EKHKHQYKIVEERDELGLYKNRBOEQSSAKVALETLSNIRNELVSLKKQLEVEKEEK 745
Qy 473 ----AVQDHQIVTVSRKSOE 488
Db 746 EKLMEQENTAILTKDKKK 765

RESULT 11
Tl6270
hypothetical protein F35D11.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: Tl6270
R:Fulton, B.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F35D11.
A:Reference number: Zl8487
A:Accession: Tl6270
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1827 <FUL>
A:Cross-references: UNIPROT:Q20042; EMBL:U29381; NID:9868214; PID:9868224; PIDN:AA68757
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F35D11.11
A:Introns: 76/2; 131/3; 159/3; 185/3; 221/3; 253/3; 320/1; 869/3; 1133/3; 1205/2; 1350/1

Query Match 9.2%; Score 239.5; DB 2; Length 1827;
Best Local Similarity 24.7%; Pred. No. 0.0026;
Matches 123; Conservative 96; Mismatches 177; Indels 101; Gaps 24;

Qy 42 KALELKNQTLRA---DEILPSESQKQDYESSWDSSELCT-----VSQKDV 86
Db 448 KEREDSHEALRSKDTTEERSRFEKERSERKIRRESDELRETQRSLSKGDVAMKTDLDK 507
Qy 87 CLPKATHQKEIDKINGKLE---ESPNDNGFLKPCRMKVSIPTKALELMDMOTFKAPPE 143
Db 508 TLRLDMLTEDEKLKKLETETEQADQDLIAECRAK-----LDEMAEKEALRK 558
Qy 144 KPSAFPAIEMQ---KSPVKNALKEQLTADRADEILPSSKQKQVENSWSDESLEET 199
Db 559 ELAEFOAIITAMEGEGKLNOBQFLESKNELNTLQI---ESLNSEVENKNEIRNLMAT 615
Qy 200 VSQKDVCPVK---ATHQ---KEMDKISGKLED-STLSKILDTVHSCEARLEQK-----247
Db 616 LOKEKVHIQNVRTSSHQLTATYVEEANGIDILKAEITLHEQVN--ERTROISEANEKYD 673
Qy 248 DHCEORTGKMEQ---KKKFCVLKKLSEAKEIKSOLENKVQKWEQLCSVR-----LTL 299
Db 674 DAARKNDALLEDVATWQEKVEQLKMELEEMNRRGOEKEREADLRLALDDLRGNFKLT-732
Qy 300 NOEESKRRNADILNEKI-----REELGRIEE-QHRKELEVKQOOLEQALRIQ 344
Db 733 NELKQGVTVDSLNLEISLSLKEQLNKSKEKREELLRMEELEQKNEAEKVEYEVKLQILA 792

```

Qy 345 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLTKHQYQEKENKYFDI 404
Db 793 EKDRQGV-NFGKECEARMTLTKEH--MLMEEDQLKVD-----HLHTE-----BEV 838
Qy 405 KILKEK-NAELQMTLKLKEESTKTRASQVSGQLKVLIAENTMLT-----SKLKEK 453
Db 839 ERLKEKMRKELE---KLNEQNDGDAEWSNERNRLSSKNEAVTELQERVOKLEDVVKKE 895
Qy 454 QDKKEI-LEAEIESHHPR 469
Db 896 EDKEIALRDLDSHEK 912

RESULT 12
S67593
transport protein USO1 - Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D2552; protein YDL058W
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C:Accession: S67593; A38455; S30782
R:Bioecker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67587
A:Accession: S67593
A:Molecule type: DNA
A:Residues: 1-1790 <BLO>
A:Cross-references: UNIPROT:Q07380; EMBL:Z74106; NID:G1431058; PID:G1431059
A:Experimental source: strain S288C
R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamaeaki, M.
J. Cell Biol. 113, 245-260, 1991
A:Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transp
A:Reference number: A38455; MUID:91185402; PMID:2010462
A:Accession: A38455
A:Molecule type: DNA
A:Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>
A:Cross-references: GB:X43478; NID:G94777; PID:CAA38253.1; PID:G94778
A:Note: the authors translated the codon ACT for residue 768 as Ile
R:Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
submitted to the EMBL Data Library, February 1993
A:Description: An integrin analogue in Saccharomyces cerevisiae.
A:Reference number: S30782
A:Accession: S30782
A:Molecule type: DNA
A:Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580, 'S'
A:Cross-references: EMBL:L03188
C:Genetics:
A:Gene: SGD:USO1; INT1
A:Cross-references: SGD:S0002216; MIPS:YDL058W
A:Map position: 4L
C:Keywords: coiled coil; transmembrane protein
F:326-342/Domain: transmembrane #status predicted <TM1>
F:394-410/Domain: transmembrane #status predicted <TM2>
F:617-633/Domain: transmembrane #status predicted <TM3>

Query Match 9.2%; Score 239; DB 2; Length 1790;
Best Local Similarity 22.6%; Pred. No. 0.0026;
Matches 134; Conservative 102; Mismatches 218; Indels 138; Gaps 23;

Qy 4 SIPT--KALELMDMTQTKAEP---PEKPSAFEPAIEHQKSPVNPKALELKNQETLRADEIL 58
Db 840 SIKTLEKLEITLSQKKAEDGINKMGKDLFALSGEMQAVEEN----- 882
Qy 59 PSEKQKDYESSWDSSESLCTVSOQVCLPKATHQKEIDKINGKLEESPDNDGFLKAPC 118
Db 883 -CKNLQEKDKSNVNHQK--ETKSLKEDIKAIT--EIKAINENLEE-----MKIQ 929
Qy 119 ----RMKVSIPITKALELMDMTQTKAEPPEKPSAFEPAIEHQKSPVNP--KALELKNQETLR 172
Db 930 NNLKKEKEHISKELVE-----YKSRFQSHDLNVLAKTEKLKSLANNYKDMQAEENSILK 983
Qy 173 ADQWFPSEK-----QKKVEENSWDSSE-----LRETVSQ-----KDVCPV 208
Db 984 AVEESKNESIQLSNLQNKIDMSQOEKENFOIERGSIEFKNTEQLKKTITSDLEQTEKIIIS 1043

```

```

Qy 209 KA-----THQKMDKISGKLEDSTS-----LSKILDTVHSCERARELOKDHCEQRTG 255
Db 1044 KSDSSKDEYBSQISLLKELKLTATTANDENVNKISELTKTREELAEALAAVKNLKVELET 1103
Qy 256 KMEQMKKFCVVLKKLSEAKKIKSOLENQKVWQELCSVRLTLNQBSEKERNADILNEK 315
Db 1104 KLETSEKALKVEKNEEHLKEEKIQLEKEATETKQOLNSLRANLESLEKHEDLAAQILKK 1163
Qy 316 IREELGRIEEOHRKELE-----VKQLEQALRIQDIELKSVESNLNQVSHTHENYVL 368
Db 1164 YEQIANKERYNEEISQLNDEITSTQSNESIKKQDEL---EGEVAMKSTSEQNL 1220
Qy 369 LHENCMLKKEIAMLKLEIATLTKHQYQEKENKYFDIKIL-----KEKN- 411
Db 1221 ----KSEIDALNLQIKELKKNETNEASLLSSIKSVESETVKIKELQDECNPKKEV 1274
Qy 412 AELQMTLKLKE-----ESLTKRASQVSGQLKVLIAENTMLTSLKKEQDKKEI 458
Db 1275 SELEDKLKASDKNSKYLELOKSEKIKKEELDAAKTTELKIQLEKITNL-SKAKEKSESEL 1333
Qy 459 LEAEIESHHPRLASAVODHDOIVTSRKSOEPAFIAGDACLQKRMNVDSVT 510
Db 1334 SRLKTSSEER-KNABEQLEKLNKIEIQKQAF-----KERLLNEGST 1378

RESULT 13
A56539
giantin - human
N:Alternate names: macrogolin
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 26-Jan-1996 #text_change 09-Jul-2004
C:Accession: A56539; S37536
R:Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M.
Mol. Cell. Biol. 14, 2564-2576, 1994
A:Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein (
A:Reference number: A56539; MUID:94187728; PMID:7511208
A:Accession: A56539
A:Molecule type: mRNA
A:Residues: 1-3259 <SEE>
A:Cross-references: UNIPROT:Q14789; EMBL:X75304; NID:G405714; PID:CAA53052.1; PID:G4057
C:Genetics:
A:Gene: GDB:GOLGB1; GCP; GCP371
A:Cross-references: GDB:454958
A:Map position: 3q13.31-3q13.31
C:Superfamily: giantin
C:Keywords: coiled coil; Golgi apparatus; transmembrane protein
F:3238-3254/Domain: transmembrane #status predicted <TM>

```

```

Query Match 9.2%; Score 239; DB 1; Length 3259;
Best Local Similarity 19.5%; Pred. No. 0.0051;
Matches 114; Conservative 116; Mismatches 235; Indels 120; Gaps 17;

Qy 8 KALELMDMTQTKAEPPEKPSAFEPAIEHQKSPVNPKALEL-----KNEQTLRADEILPS 60
Db 1594 KSSKIAESTEWQEKHEKLEQYELLQSYENVSNESAEIRIQHVBAVRQEKQELYGKLKST 1653
Qy 61 ESKQKDYESSWDSSESLCTVSOQVCLPKATHQK--EIDKINGKLEESPDNDGFLKAPC 118
Db 1654 EANKKETEKOLQEAQEBQMEEMKEMKRPKSKQKQKILEBENDRLRAEVPAGDTAKEC 1713
Qy 119 RMKVSIPITKALELMDMTQTKAEPPEKPSAFEPAIEHQKSPVNPKALELKNQETLRADQWFP 178
Db 1714 -METLLSSNASKKEELERVKMEYETLSKKFQSLMSKDSLSEEVQDLKH-----QLED 1765
Qy 179 SESQKQKVEENSWDSSESLRETIVSOQVCPKATHQKMDKISGKLESTSLSKILDTVHS 238
Db 1766 NVSRQANLEATEKHNDQTNVT-----EEGTQSIPTGETEEDSLS--MSTRPT 1810
Qy 239 CERARELOK-----DHCEORTGMEQMKKFCVLKKLSEAKKIKSOLENQK 285
Db 1811 CSESVPKSAKSNPAVKDFSSHDSINNLYQIQIDQLKRIAGLEBEKQNKKEFSQTLNEK 1870

```


Search completed: February 19, 2005, 02:12:05
Job time : 18.4804 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 02:11:01 ; Search time 48.0894 Seconds
(without alignments)
3484.083 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDQTFKA.....IAGDACLQRKXNDVDSSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2599	100.0	512	16	US-10-181-663-16
2	2589	99.6	1341	13	US-10-007-805-565
3	2589	99.6	1341	14	US-10-076-622-565
4	2589	99.6	1341	14	US-10-177-293-334
5	2589	99.6	1341	14	US-10-124-805-565
6	2589	99.6	1341	15	US-10-088-270A-4
7	2589	99.6	1349	13	US-10-007-805-573
8	2589	99.6	1349	14	US-10-076-622-573
9	2589	99.6	1349	14	US-10-124-805-573
10	2578	99.2	1002	9	US-09-604-287A-475
11	2578	99.2	1002	9	US-09-834-759-475
12	2578	99.2	1002	10	US-09-551-621-475
13	2578	99.2	1002	13	US-10-007-805-475

14	2578	99.2	1002	14	US-10-076-622-475	Sequence 475, App
15	2578	99.2	1002	14	US-10-124-805-475	Sequence 475, App
16	2578	99.2	1095	9	US-09-834-759-493	Sequence 493, App
17	2578	99.2	1095	13	US-10-007-805-493	Sequence 493, App
18	2578	99.2	1095	14	US-10-076-622-493	Sequence 493, App
19	2578	99.2	1095	14	US-10-124-805-493	Sequence 493, App
20	2576	99.1	1013	13	US-10-007-805-553	Sequence 553, App
21	2576	99.1	1013	14	US-10-076-622-553	Sequence 553, App
22	2576	99.1	1013	14	US-10-124-805-553	Sequence 553, App
23	2125	81.8	445	9	US-09-825-301-29	Sequence 29, Appl
24	2125	81.8	445	9	US-09-604-287A-473	Sequence 473, App
25	2125	81.8	445	9	US-09-834-759-473	Sequence 473, App
26	2125	81.8	445	10	US-09-551-621-473	Sequence 473, App
27	2125	81.8	445	13	US-10-007-805-473	Sequence 473, App
28	2125	81.8	445	14	US-10-076-622-473	Sequence 473, App
29	2125	81.8	445	14	US-10-124-805-473	Sequence 473, Appl
30	2125	81.8	445	14	US-10-033-527-29	Sequence 29, Appl
31	1781	68.5	1225	14	US-10-177-293-332	Sequence 332, App
32	1775	68.3	1239	13	US-10-007-805-577	Sequence 577, App
33	1775	68.3	1239	14	US-10-076-622-577	Sequence 577, App
34	1775	68.3	1239	14	US-10-124-805-577	Sequence 577, App
35	1542	59.3	1011	16	US-10-408-765A-1557	Sequence 1557, App
36	1513.5	58.2	650	9	US-09-825-301-25	Sequence 25, Appl
37	1513.5	58.2	650	9	US-09-604-287A-469	Sequence 469, App
38	1513.5	58.2	650	9	US-09-834-759-469	Sequence 469, App
39	1513.5	58.2	650	10	US-09-551-621-469	Sequence 469, App
40	1513.5	58.2	650	13	US-10-007-805-469	Sequence 469, App
41	1513.5	58.2	650	14	US-10-076-622-469	Sequence 469, App
42	1513.5	58.2	650	14	US-10-124-805-469	Sequence 469, App
43	1513.5	58.2	650	14	US-10-033-527-25	Sequence 25, Appl
44	1513.5	58.2	743	9	US-09-834-759-494	Sequence 494, App
45	1513.5	58.2	743	13	US-10-007-805-494	Sequence 494, App

ALIGNMENTS

RESULT 1

US-10-181-663-16
; Sequence 16, Application US/10181663
; Publication No. US20040126398A1
; GENERAL INFORMATION:
; APPLICANT: Jager, Dirk
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gure, Ali
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd
; APPLICANT: Chen, Yao-tseng
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated Antigens per se, and Uses Thereof
; FILE REFERENCE: LUD 5615
; CURRENT APPLICATION NUMBER: US/10/181.663
; CURRENT FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 09/451,739
; PRIOR FILING DATE: 1999-11-30
; NUMBER OF SEQ ID NOS: 19
; SEQ ID NO 16
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-181-663-16

Query Match 100.0%; Score 2599; DB 16; Length 512;
Best Local Similarity 100.0%; Pred. No. 1.7e-141;
Matches 512; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDQTFKAEPKPSAFPAIEQMKSVPNKALELKNQETLRADEILPS 60
Db 1 MKVSIPTKALELMDQTFKAEPKPSAFPAIEQMKSVPNKALELKNQETLRADEILPS 60
Qy 61 ESKQDYEESSWDSSESICETVSQKDVCLPKATHOKEIDKINGKLEESPDNGFLKAPCRM 120

Db 61 ESKQKDYESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180
Db 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180
Qy 181 SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC 240
Db 181 SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC 240
Qy 241 RARELOKHCEORTGMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRLTLN 300
Db 241 RARELOKHCEORTGMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRLTLN 300
Qy 301 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSVEENLNQVSH 360
Db 301 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSVEENLNQVSH 360
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Qy 481 VTSRKSQEPAFHAGDACLQKRNVDVSTDI 512
Db 481 VTSRKSQEPAFHAGDACLQKRNVDVSTDI 512

RESULT 2
US-10-007-805-565
; Sequence 565, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C10
; CURRENT FILING DATE: 2001-12-07
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-565
Query Match 99.6%; Score 2589; DB 13; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADILPS 60
Db 677 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADILPS 736
Qy 61 ESKQKDYESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQKDYESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180
Db 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180

Db 797 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 856
Qy 181 SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC 240
Db 857 SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC 916
Qy 241 RARELOKHCEORTGMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRLTLN 300
Db 917 RARELOKHCEORTGMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSVEENLNQVSH 360
Db 977 QEEKRRNADILNEKIREELGRIEBOHRKELEVQKQLEQALRIODIELKSVEENLNQVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 1156
Qy 481 VTSRKSQEPAFHAGDACLQKRNVDVSTDI 510
Db 1157 VTSRKSQEPAFHAGDACLQKRNVDVSTDI 1186

RESULT 3
US-10-076-622-565
; Sequence 565, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-565
Query Match 99.6%; Score 2589; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADILPS 60
Db 677 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADILPS 736
Qy 61 ESKQKDYESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQKDYESSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180
Db 797 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 856
Qy 181 SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC 240
Db 857 SKQKVEENSWDSLSLCEVTSQKDVCLPKATHQKEMDKISGKLESDSTLSKILDTVHSC 916
Qy 241 RARELOKHCEORTGMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRLTLN 300
Db 917 RARELOKHCEORTGMEQMKKFCVLKKLSEAKEIKSOLENOKVWEQELCSVRLTLN 976

QY 301 QEEKRRNADILNEKIREELGRIBEQHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEQHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 1036
QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 1096
QY 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 1156
QY 481 VTSRKSQEPAFHIAAGDACLQRKMNVDVSS 510
Db 1157 VTSRKSQEPAFHIAAGDACLQRKMNVDVSS 1186

RESULT 4

US-10-177-293-334
; Sequence 334, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Puzstai, Lajos
; APPLICANT: Mexic, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 334
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-334

Query Match 99.6%; Score 2589; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 60
Db 677 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 736

QY 61 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
QY 121 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 180
Db 797 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 856
QY 181 SKQKVEENSWDSSESLRETQVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
Db 857 SKQKVEENSWDSSESLRETQVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 916
QY 241 RARELQDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLEQENQKVKWEQELCSVRLTLN 300
Db 917 RARELQDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLEQENQKVKWEQELCSVRLTLN 976
QY 301 QEEKRRNADILNEKIREELGRIBEQHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEQHREKLEVKQOQLRQALRIQDIELKSVESNLNQVSH 1036
QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL 1096
QY 421 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGOLKVLIAENTMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQI 1156
QY 481 VTSRKSQEPAFHIAAGDACLQRKMNVDVSS 510
Db 1157 VTSRKSQEPAFHIAAGDACLQRKMNVDVSS 1186

RESULT 5

US-10-124-805-565
; Sequence 565, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-124-805-565

Query Match 99.6%; Score 2589; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 60
Db 677 MKVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 736
QY 61 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQDYEESSWDSSESLCETVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
QY 121 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 180
Db 797 KVSIPTKALELMDMQTFKAEPPEKPSAFEPALIEQKSVPNKALELKNQETLRADQIFLPS 856
QY 181 SKQKVEENSWDSSESLRETQVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
Db 857 SKQKVEENSWDSSESLRETQVSKQDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 916

Qy 241 RARELOKHCHORTGMEQMKKFCVLKKLSEAKETKSOLENKVWEQELCSVRLTLN 300
Db 917 RARELOKHCHORTGMEQMKKFCVLKKLSEAKETKSOLENKVWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNQVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNVLAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNVLAELQMTLKL 1096
Qy 421 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKQDKEILEAEIESHPRLASAVODHDOI 480
Db 1097 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKQDKEILEAEIESHPRLASAVODHDOI 1156
Qy 481 VTSRKSQEPAPHIAGDACLQRMNVDSST 510
Db 1157 VTSRKSQEPAPHIAGDACLQRMNVDSST 1186

RESULT 6

US-10-058-270A-4

; Sequence 4, Application US/10058270A

; Publication No. US20040029114A1

; GENERAL INFORMATION:

; APPLICANT: Mack, David H.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Afar, Daniel

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and Methods of Screening for Modulators of Breast Cancer

; FILE REFERENCE: 018501-005210US

; CURRENT APPLICATION NUMBER: US/10/058,270A

; CURRENT FILING DATE: 2002-01-24

; PRIOR APPLICATION NUMBER: US 60/263,965

; PRIOR FILING DATE: 2001-01-24

; PRIOR FILING DATE: 2001-02-02

; PRIOR FILING DATE: 2001-04-09

; PRIOR FILING DATE: 2001-04-09

; PRIOR FILING DATE: 2001-05-04

; PRIOR FILING DATE: 2001-05-29

; NUMBER OF SEQ ID NOS: 141

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1341

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-058-270A-4

Query Match 99.6%; Score 2589; DB 15; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.9e-140; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0;

Qy 1 MKVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 60
Db 677 MKVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 736
Qy 61 ESKQDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 737 ESKQDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADQMPFSE 180
Db 797 KVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADQMPFSE 856
Qy 181 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 240

Db 857 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 916
Qy 241 RARELOKHCHORTGMEQMKKFCVLKKLSEAKETKSOLENKVWEQELCSVRLTLN 300
Db 917 RARELOKHCHORTGMEQMKKFCVLKKLSEAKETKSOLENKVWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 977 QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNQVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNVLAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNVLAELQMTLKL 1096
Qy 421 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKQDKEILEAEIESHPRLASAVODHDOI 480
Db 1097 KEESLTTRASQYSGQLKVLIAENTMTLSKLEKQDKEILEAEIESHPRLASAVODHDOI 1156
Qy 481 VTSRKSQEPAPHIAGDACLQRMNVDSST 510
Db 1157 VTSRKSQEPAPHIAGDACLQRMNVDSST 1186

RESULT 7

US-10-007-805-573

; Sequence 573, Application US/10007805

; Publication No. US20020150581A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yugu

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Durham, Margarita

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.470C10

; CURRENT APPLICATION NUMBER: US/10/007,805

; CURRENT FILING DATE: 2001-12-07

; NUMBER OF SEQ ID NOS: 593

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 573

; LENGTH: 1349

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-007-805-573

Query Match 99.6%; Score 2589; DB 13; Length 1349;
Best Local Similarity 100.0%; Pred. No. 1.9e-140; Indels 0; Gaps 0;
Matches 510; Conservative 0; Mismatches 0;

Qy 1 MKVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 60
Db 695 MKVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADIELPS 744
Qy 61 ESKQDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 745 ESKQDYESSWDSSESLCETVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 804
Qy 121 KVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADQMPFSE 180
Db 805 KVSIPTKALELMDMQTFKAEPPKPSAFEPAPAEIMQKSVPNKALELKNQETLRADQMPFSE 864
Qy 181 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 240
Db 865 SKQKKEVNSWDSSESLRETQVTSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDTVHSCE 924

Qy	241	RARELOXDHC	EQRTGKWEOMKKFCVILKKKLSEAKEIKS	QLENOKVWEQBELCSVR	TLN	300
Db	925	RARELOXDHC	EQRTGKWEOMKKFCVILKKKLSEAKEIKS	QLENOKVWEQBELCSVR	TLN	984
Qy	301	QEEKRNADILNE	KIREELGRIBEOHRKELEVQOQLEQALRIQD	IELKSYESNLNVSH		360
Db	985	QEEKRNADILNE	KIREELGRIBEOHRKELEVQOQLEQALRIQD	IELKSYESNLNVSH		1044
Qy	361	THENENYLLHNC	MLKKEIAMLKLEIATLKHQYOEKENKYPED	IKILKEKNAELQMTLKL		420
Db	1045	THENENYLLHNC	MLKKEIAMLKLEIATLKHQYOEKENKYPED	IKILKEKNAELQMTLKL		1104
Qy	421	KEESLTKRASQ	YSQGLKVLTAEINTMLTSLKEKQDKEILEAEIESHHPRLASAVODHDOI			480
Db	1105	KEESLTKRASQ	YSQGLKVLTAEINTMLTSLKEKQDKEILEAEIESHHPRLASAVODHDOI			1164
Qy	481	VTGRKQEPAPH	IAGDACLQRKNMVDVSST			510
Db	1165	VTGRKQEPAPH	IAGDACLQRKNMVDVSST			1194

RESULT 8

US-10-076-622-573

; Sequence 573, Application US/10076622

; Publication No. US20030023036A1

GENERAL INFORMATION:

APPLICANT: Houghton, Raymond L.

APPLICANT: Sleath, Paul R.

APPLICANT: Persing, David H.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C11

; CURRENT APPLICATION NUMBER: US/10/076,622

; CURRENT FILING DATE: 2002-02-13

; NUMBER OF SEQ ID NOS: 627

```
; SOFTWARE: FastSEQ for Windows Version 4.0
```

; SEQ ID NO 573

; LENGTH: 1349

```

; TYPE: PRT

```

; ORGANISM: Homo sapiens

US-10-076-622-573

Query Match 99.6%; Score 2589; DB 14; Length 1349;

Best Local Similarity 100.0%; Pred. No. 1.9e-140;

Matches	510:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps	0:
----------------	-------------	---------------------	-----------	-------------------	-----------	---------------	-----------	-------------	-----------

Qy	1	MKVSIPTKALELMDMQTFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQOTLRADEILPS	60
Db	685	MKVSIPTKALELMDMQTFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQOTLRADEILPS	744
Qy	61	ESKQKDYEESSWDSSESJCETVSQKDVCCLKPATHQKEIDKINGKLEESPDDNDGFLKAPCRM	120
Db	745	ESKQKDYEESSWDSSESJCETVSQKDVCCLKPATHQKEIDKINGKLEESPDDNDGFLKAPCRM	804
Qy	121	KVSIPTKALELMDMQTFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQOTLRAQOMFPSE	180
Db	805	KVSIPTKALELMDMQTFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQOTLRAQOMFPSE	864
Qy	181	SKQKVVENSWSDESJLRETQVSKDVCVPKATHQKEMDKISGLKBDSTSLSKILDTVHSCE	240
Db	865	SKQKVVENSWSDESJLRETQVSKDVCVPKATHQKEMDKISGLKBDSTSLSKILDTVHSCE	924
Qy	241	RARELQKDHCBOETQKMEOMKKFCVLKKLSEAKEIKSQLENOKVMEQELCSVRJTLN	300
Db	925	RARELQKDHCBOETQKMEOMKKFCVLKKLSEAKEIKSQLENOKVMEQELCSVRJTLN	984
Qy	301	QEBEKRNRADILNEKIREELGRIBEOHRKELEVQQLFEQALRIODIELKSVESNLQVSH	360
Db	985	QEBEKRNRADILNEKIREELGRIBEOHRKELEVQQLFEQALRIODIELKSVESNLQVSH	1044
Qy	361	THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILUKENNAELQMTLKL	420
Db	1045	THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILUKENNAELQMTLKL	1104

RESULT 10

US-09-604-287A-475

```

US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent NO. US20020064872A1
; GENERAL INFORMATION:
;
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-475

```

Query Match	99.2%;	Score 2578;	DB 9;	Length 1002;
Best Local Similarity	99.6%;	Pred. No. 5.8e-140;		
Matches 508;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	MKVSIPTKALELMDMOTFKABPPPKPSAFEPAL	EMQKSVPNKALELKNQETLRADEILPS	60
Db	338	MKVSIPTKALELMDMOTFKABPPPKPSAFEPAL	EMQKSVPNKALELKNQETLRADEILPS	397
Qy	61	ESKQKDYESSWDSSESLSCTVSKQDVCLPKATHQ	KEIDKINGKLEESPDNDGFLKAPCRM	120
Db	398	ESKQKDYESSWDSSESLSCTVSKQDVCLPKATHQ	KEIDKINGKLEESPDNDGFLKAPCRM	457
Qy	121	KVSIPTKALELMDMOTFKABPPPKPSAFEPAL	EMQKSVPNKALELKNQETLRADOMPFSE	180
Db	458	KVSIPTKALELMDMOTFKABPPPKPSAFEPAL	EMQKSVPNKALELKNQETLRADOMPFSE	517
Qy	181	SKQKKVENSNDSESLSRTVSKQDVCPKATHQEMDK	ISGKLEDTSLSKILDIVHSCE	240
Db	518	SKQKKVENSNDSESLSRTVSKQDVCPKATHQEMDK	ISGKLEDTSLSKILDIVHSCE	577
Qy	241	RARELQKDCHORTGKMEQMKKKFCVLKKKLS	SEAKEIKSQLENQKVKVEQELCSVRLTLN	300
Db	578	RARELQKDCHORTGKMEQMKKKFCVLKKKLS	SEAKEIKSQLENQKVKVEQELCSVRLTLN	637
Qy	301	QEEKRRNADILNEKIREELGRIEBQHRKELEV	KQOLFQALRIQIDIELKSVESNLNQVSH	360
Db	638	QEEKRRNADILNEKIREELGRIEBQHRKELEV	KQOLFQALRIQIDIELKSVESNLNQVSH	697
Qy	361	THENENYLLHNCMLKKEIAMLKLEIATLKHOYQ	EKENKYPEDIKILKEKNAELQMTLKL	420
Db	698	THENENYLLHNCMLKKEIAMLKLEIATLKHOYQ	EKENKYPEDIKILKEKNAELQMTLKL	757
Qy	421	KEESLTKRASQYSGQLKVLIAENTWLTSLKLEK	QDKELIEABIESHHPRLASAVQDHQI	480
Db	758	KEESLTKRASQYSGQLKVLIAENTWLTSLKLEK	QDKELIEABIESHHPRLASAVQDHQI	817
Qy	481	VTSRKSQBPAFHIAAGDACLQRKNQVDSST	510	
Db	818	VTSRKSQBPAFHIAAGDACLQRKNQVDSST	847	

RESULT 11
US-09-834-759-475
; Sequence 475, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:

```

; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocke, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PR1
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1) .. (1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-834-759-475

```

Query Match	99.2%	Score 2578;	DB 9;	Length 1002;
Best Local Similarity	99.6%;	Pred. No. 5.8e-140;		
Matches 508;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
Qy	1	MKVSIPTKALELMDMOTFFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQEOTLRADEILPS	60	
Db	338	MKVSIPTKALELMDMOTFFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQEOTLRADEILPS	397	
Qy	61	ESKQKQYEESWDSSESILCEIVSQDVCLPRATHQKEIDKINGKLEESPDNDGFLKAPCRM	120	
Db	398	ESKQKQYEESWDSSESILCEIVSQDVCLPRAXHQKEIDKINGKLEESPDNDGFLKAPCRM	457	
Qy	121	KVSIPTKALSLMDMOTFFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQEOTLRADOMPPE	180	
Db	458	KVSIPTKALSLMDMOTFFKAEPPEKPSAFAPAIEMQKSVPNKALELKNQEOTLRADOMPPE	517	
Qy	181	SKQKQYEENSWDSSESILRETYSQDVCPKATHQKEMDKISGKLEDSLSLSKILDTVHSCE	240	
Db	518	SKQKQYEENSWDSSESILRETYSQDVCPKATHQKEMDKISGKLEDSLSLSKILDTVHSCE	577	
Qy	241	RARELQKDHCEQRTGMEQMKKFCVLKXKLSEAKEIKSQLENOKVWQEQLCSVRLTLN	300	
Db	578	RARELQKDHCEQRTGMEQMKKFCVLKXKLSEAKEIKSQLENOKVWQEQLCSVRLTLN	637	
Qy	301	QSEKRRNADILNEKIREELGRIBEEQHRKELEVQQLQEQALRIODIELKSVESNLNQVSH	360	
Db	638	QSEKRRNADILNEKIREELGRIBEEQHRKELEVQQLQEQALRIODIELKSVESNLNQVSH	697	
Qy	361	THENENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENKYPEDIKILKEKNAEIQMTLKL	420	
Db	698	THENENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENKYPEDIKILKEKNAEIQMTLKL	757	
Qy	421	KEESLTKRASQYSGQLKVLIAENTMLTSLKKEKQDKKEILEAEIESHHPRLASAVQDHDQI	480	
Db	758	KEESLTKRASQYSGQLKVLIAENTMLTSLKKEKQDKKEILEAEIESHHPRLASAVQDHDQI	817	
Qy	481	VTSKSQEPAFHIAAGDACLQRMNVVDSST	510	
Db	818	VTSKSQEPAFHIAAGDACLQRMNVVDSST	847	

RESULT 12
US-09-551-621-475
; Sequence 475, Application US/09551621
; Publication No. US20030104366A1
; GENERAL INFORMATION: Jiang
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Miccham, Jennifer L.

; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2006-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Query Match 99.2%; Score 2578; DB 10; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 180
Db 458 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 517
Qy 181 SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEEDSTSLKILDTVHSC 240
Db 518 SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEEDSTSLKILDTVHSC 577
Qy 241 RARELQDHCORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 300
Db 578 RARELQDHCORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 637
Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSVESNLNVSH 360
Db 638 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSVESNLNVSH 697
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQVSGQLKVLIAENTMLTSKLEKQDKTELEAIEISHHPRLASAVQDDHDI 480
Db 758 KEESLTKRASQVSGQLKVLIAENTMLTSKLEKQDKTELEAIEISHHPRLASAVQDDHDI 817
Qy 481 VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 510
Db 818 VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 847

RESULT 13

US-10-007-805-475
; Sequence 475, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-007-805-475

Query Match 99.2%; Score 2578; DB 13; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 180
Db 458 KVSIPTKALELMDMOTFKAEPPKPSAFEPALIEQKSVPNKALELKNQETLRADQMPFSE 517
Qy 181 SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEEDSTSLKILDTVHSC 240
Db 518 SKQKVVENSWSLSRLTQKDVCPKATHQKEMDKISGKLEEDSTSLKILDTVHSC 577
Qy 241 RARELQDHCORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 300
Db 578 RARELQDHCORTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQLCSVRLTLN 637
Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSVESNLNVSH 360
Db 638 QEEKRRNADILNEKIREELGRIEBQHRKELEVQKQLEALRIQDIELKSVESNLNVSH 697
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQVSGQLKVLIAENTMLTSKLEKQDKTELEAIEISHHPRLASAVQDDHDI 480
Db 758 KEESLTKRASQVSGQLKVLIAENTMLTSKLEKQDKTELEAIEISHHPRLASAVQDDHDI 817
Qy 481 VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 510
Db 818 VTSRKSQEPAPHIAGDAQLQRKMNVDVSSST 847

RESULT 14

US-10-076-622-475
; Sequence 475, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076.622

; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-124-805-475

Query Match 99.2%; Score 2578; DB 14; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 397

Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457

Qy 121 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 180
Db 458 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 517

Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
Db 518 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 577

Qy 241 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKEQELCSVRLTLN 300
Db 578 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKEQELCSVRLTLN 637

Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVKQOQLRQALRIQDIELKSVESNLQVSH 360
Db 638 QEEKRRNADILNEKIREELGRIEBQHRKELEVKQOQLRQALRIQDIELKSVESNLQVSH 697

Qy 361 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757

Qy 421 KEESLTRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 480
Db 758 KEESLTRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 817

Qy 481 VTSRKQEPAFHIAGDACLQKMNVDVSSST 510
Db 818 VTSRKQEPAFHIAGDACLQKMNVDVSSST 847

Search completed: February 19, 2005, 02:33:41
Job time : 50.0894 secs

; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-076-622-475

Query Match 99.2%; Score 2578; DB 14; Length 1002;
Best Local Similarity 99.6%; Pred. No. 5.8e-140;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 60
Db 338 MKVSIPTKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 397

Qy 61 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457

Qy 121 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 180
Db 458 KVSIPITKALELMDMOTFKAEPPKPSAFEPALAEIEMQKSVPNKALELKNQETLRADILPS 517

Qy 181 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 240
Db 518 SKQKVEENSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 577

Qy 241 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKEQELCSVRLTLN 300
Db 578 RARELQKHCEQRTGMEQMKKFCVLKKLSEAKEIKSQLENQKVKEQELCSVRLTLN 637

Qy 301 QEEKRRNADILNEKIREELGRIEBQHRKELEVKQOQLRQALRIQDIELKSVESNLQVSH 360
Db 638 QEEKRRNADILNEKIREELGRIEBQHRKELEVKQOQLRQALRIQDIELKSVESNLQVSH 697

Qy 361 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 698 THENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757

Qy 421 KEESLTRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 480
Db 758 KEESLTRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHQI 817

Qy 481 VTSRKQEPAFHIAGDACLQKMNVDVSSST 510
Db 818 VTSRKQEPAFHIAGDACLQKMNVDVSSST 847

RESULT 15
US-10-124-805-475
; Sequence 475, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:52:33 ; Search time 55.7765 Seconds
(without alignments)
3550.262 Million cell updates/sec

Title: US-09-602-362E-16

Perfect score: 2599

Sequence: 1 MKVSIPTKALELMDMQTFKA.....IAGDACLQRKNVDVSTDI 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2599	100.0	512	4	AAB84701 Amino aci
2	2589	99.6	1341	4	AAB84702 Amino aci
3	2589	99.6	1341	5	ABJ05537 Breast ca
4	2589	99.6	1341	6	ABJ37784 Human tum
5	2589	99.6	1341	6	ABJ37784 Human tum
6	2589	99.6	1341	7	ABJ37784 Human tum
7	2589	99.6	1341	7	ADL93227 Human bre
8	2589	99.6	1349	7	ADL93227 Human bre
9	2578	99.2	1002	4	AU333351 Human bre
10	2578	99.2	1002	5	ABG78918 Human bre
11	2578	99.2	1002	6	ABJ37741 Human tum
12	2578	99.2	1002	7	ADL93137 Human tum
13	2578	99.2	1002	8	ADL93137 Human tum
14	2578	99.2	1095	4	AU333357 Human bre
15	2578	99.2	1095	5	ABG78924 Human bre
16	2578	99.2	1095	6	ABJ37747 Human tum
17	2578	99.2	1095	7	ADL93155 Human tum
18	2576	99.1	1013	6	ABJ37783 Human tum
19	2576	99.1	1013	7	ADL93215 Human bre
20	2125	81.8	445	4	AAB50249 Human bre
21	2125	81.8	445	4	AAG65987 B726P spl
22	2125	81.8	445	4	AU333350 Human bre
23	2125	81.8	445	5	ABG78917 Human bre
24	2125	81.8	445	6	ABJ37740 Human tum
25	2125	81.8	445	7	ADL93135 Human bre

ALIGNMENTS

RESULT 1

AAB84701
ID AAB84701 standard; protein; 512 AA.

XX AC AAB84701;

XX XX 17-SEP-2001 (first entry)

XX DE Amino acid sequence of a human cancer associated antigen.

XX XX Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.

XX OS Homo sapiens.

XX PN WO200147959-A2.

XX PD 05-JUL-2001.

XX PF 29-NOV-2000; 2000WO-US042334.

XX PR 30-NOV-1999; 99US-00451739.

XX PR 24-OCT-2000; 2000US-00602362.

XX PA (LUDW-) LUDWIG INST CANCER RES.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX (CORR) CORNELL RES FOUND INC.

XX PI Jager D, Stockert B, Scanlan M, Knuth A, Old L, Gure A, Chen Y;

XX XX WPI; 2001-441706/47.

XX XX N-PSDB; AAH28489.

XX PT Isolated cancer associated nucleic acid molecule identified by SEREX
PT (serological identification of antigens by recombinant expression
PT cloning) technique, useful in nucleic acid based therapies to treat
cancer.

XX PS Claim 83; Page 50-51; 62pp; English.

XX CC The present sequence represents a human cancer associated antigen. The
CC sequence was identified using probes derived from the INGI gene. The INGI
CC gene is a tumour suppressor candidate gene. The cancer associated antigen
CC polynucleotides and polypeptides are useful for screening for the
CC possible presence of a pathological condition in a subject such as
CC cancer. The cancer associated antigen polypeptides are useful for
CC producing vaccines

XX SQ Sequence 512 AA;

Ade44425 Human bre
Abx47547 Breast ca
Abj37789 Human tum
Adl93239 Human bre
Aab84703 Amino aci
Adj69751 Human hea
Aab50263 Human bre
Aag65983 B726P spl
Aau33346 Human bre
Abg78913 Human bre
Abj37736 Human tum
Adl93131 Human bre
Ade44421 Human bre
Aau33358 Human bre
Abg78925 Human bre
Abj37748 Human tum
Adl93156 Human tum
Abj37782 Human bre
Adl93214 Human bre
Aab07638 Amino aci

Tue Feb 22 09:49:10 2005

us-09-602-362e-16.rag

Db 737 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMOTFKAEPEKPSAFEPALMOKSVPNKALELKNQTLRADQMPPE 180
Db 797 KVSIPTKALELMDMOTFKAEPEKPSAFEPALMOKSVPNKALELKNQTLRADQMPPE 856
Qy 181 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRM 240
Db 857 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRM 916
Qy 241 RARELQDHCCEQRTGKMEQMKKFCVLKXKLSAEIKSQLENQKVKWEQELCSVRLTLN 300
Db 917 RARELQDHCCEQRTGKMEQMKKFCVLKXKLSAEIKSQLENQKVKWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIEEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 360
Db 977 QEEKRRNADILNEKIREELGRIEEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKLEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKLEAEIESHHPRLASAVQDHDQI 1156
Qy 481 VTSRKSQBPAPHIAGDACLQKMNVDVSS 510
Db 1157 VTSRKSQBPAPHIAGDACLQKMNVDVSS 1186

RESULT 5
ABR47548
ID ABR47548 standard; protein; 1341 AA.
XX ABR47548;
AC ABR47548;
XX 12-JUN-2003 (first entry)
XX Breast cancer associated protein sequence SEQ ID NO:334.
XX Human; breast cancer; cytostatic; gene therapy.
XX Homo sapiens.
XX WO2003004989-A2.
XX 16-JAN-2003.
XX 21-JUN-2002; 2002WO-US019669.
XX 21-JUN-2001; 2001US-0299887P.
XX 27-JUN-2001; 2001US-0301572P.
XX 18-JUL-2001; 2001US-0306501P.
XX 23-SEP-2001; 2001US-0325002P.
XX 05-MAR-2002; 2002US-0362885P.
XX 14-MAY-2002; 2002US-0380391P.
XX (MILL-) MILLENIUM PHARM INC.
XX Lillie J, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S;
XX Mertens M, Monahan JE, Myer Y, Wang Y, Xu Y, Zhao X, Meyers RE;
XX Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX WPI; 2003-210381/20.
XX N-PSDB; ACC50246.
XX Breast cancer diagnosis or treatment by comparing the level of expression
XX of a marker in a patient sample with that in the control non-breast
XX cancer sample.
XX Claim 1; SEQ ID NO 334; 128pp; English.

XX The present invention describes a method for assessing whether a patient
CC is afflicted with breast cancer. The method comprises comparing the level
CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC ABR47386 to ABR47632) in a patient sample and the normal level of
CC expression of the marker in a control non-breast cancer sample, where a
CC significant increase in the level of expression of the marker in the
CC patient sample and the normal level is an indication that the patient is
CC afflicted with breast cancer. The breast cancer associated sequences from
CC the present invention have cytostatic activities and can be used in gene
CC therapy. The method is useful for diagnosing and treating breast cancer.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 1341 AA;
Query Match 99.6%; Score 2589; DB 6; Length 1341;
Best Local Similarity 100.0%; Pred. No. 3.3e-171;
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMOTFKAEPEKPSAFEPALMOKSVPNKALELKNQTLRADILPS 60
Db 677 MKVSIPTKALELMDMOTFKAEPEKPSAFEPALMOKSVPNKALELKNQTLRADILPS 736
Qy 61 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRM 120
Db 737 ESKQDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRM 796
Qy 121 KVSIPTKALELMDMOTFKAEPEKPSAFEPALMOKSVPNKALELKNQTLRADQMPPE 180
Db 797 KVSIPTKALELMDMOTFKAEPEKPSAFEPALMOKSVPNKALELKNQTLRADQMPPE 856
Qy 181 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRM 240
Db 857 SKQKVEENSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRM 916
Qy 241 RARELQDHCCEQRTGKMEQMKKFCVLKXKLSAEIKSQLENQKVKWEQELCSVRLTLN 300
Db 917 RARELQDHCCEQRTGKMEQMKKFCVLKXKLSAEIKSQLENQKVKWEQELCSVRLTLN 976
Qy 301 QEEKRRNADILNEKIREELGRIEEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 360
Db 977 QEEKRRNADILNEKIREELGRIEEOHREKLEVKQLEQALRIQDIELKSVESNLNOVSH 1036
Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 1096
Qy 421 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKLEAEIESHHPRLASAVQDHDQI 480
Db 1097 KEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKLEAEIESHHPRLASAVQDHDQI 1156
Qy 481 VTSRKSQBPAPHIAGDACLQKMNVDVSS 510
Db 1157 VTSRKSQBPAPHIAGDACLQKMNVDVSS 1186

RESULT 6
ADL93227
ID ADL93227 standard; protein; 1341 AA.
XX ADL93227;
XX 20-MAY-2004 (first entry)
XX Human breast cancer-associated polypeptide #49.
XX gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX Homo sapiens.
XX US2003166022-A1.

XX PD 04-SEP-2003.

XX XX 15-APR-2002; 2002US-00124805.

XX PF 28-DEC-1998; 98US-00222575.

XX PR 02-APR-1999; 99US-00285480.

XX PR 23-JUN-1999; 99US-00339338.

XX PR 02-SEP-1999; 99US-00389681.

XX PR 03-NOV-1999; 99US-00433826.

XX PR 17-APR-2000; 2000US-00551621.

XX PR 08-JUN-2000; 2000US-00590751.

XX PR 22-JUN-2000; 2000US-00604287.

XX PR 20-JUL-2000; 2000US-00620405.

XX PR 13-APR-2001; 2001US-00834759.

XX PR 07-DEC-2001; 2001US-00007805.

XX PR 13-FEB-2002; 2002US-0007805.

XX PR 13-FEB-2002; 2002US-00076622.

XX FA (CORI-) CORIXA CORP.

XX XX Houghton RL, Sleath PR, Persing DH;

XX XX WPI: 2003-874918/81.

XX DR N-PSDB; ADL93226.

XX XX

XX PT An isolated oncogenic polypeptide useful for preventing, diagnosing and

XX PT treating breast cancer.

XX XX

XX PS Example 12; SEQ ID NO 565; 294pp; English.

XX XX

XX CC The invention relates to an isolated breast cancer-associated

XX CC polypeptide. The polypeptide may be used for the diagnosis and treatment

XX CC of breast cancers. The methods are useful for detecting the presence of a

XX CC cancer in a patient and treating a cancer in a patient. The present

XX CC sequence represents the amino acid sequence of a human breast cancer-

XX CC associated polypeptide.

XX XX

XX SQ Sequence 1341 AA;

Query Match 99.6%; Score 2589; DB 7; Length 1341;

Best Local Similarity 100.0%; Pred. No. 3.3e-171; Indels 0; Gaps 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 60

Db 677 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 736

Qy 61 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120

Db 737 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 796

Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180

Db 797 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 856

Qy 181 SKQKVEENSWDSSELRSTVSQKDVCPKATHQKEMDKISGKLESDSTSLKILDTVHSC 240

Db 857 SKQKVEENSWDSSELRSTVSQKDVCPKATHQKEMDKISGKLESDSTSLKILDTVHSC 916

Qy 241 RARELQDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 300

Db 917 RARELQDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQBLCSVRLTLN 976

Qy 301 QEEKRRNADTLNKKIREELGRIREOHRKELEVQOOLRQIODEILKSVESNLNOVSH 360

Db 977 QEEKRRNADTLNKKIREELGRIREOHRKELEVQOOLRQIODEILKSVESNLNOVSH 1036

Qy 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 420

Db 1037 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILKEKNAELQMTLKL 1096

Qy 421 KEESLTRASQYSQGLKVLIANTWLTSLKKEQDKTEILEAIESHHPRLASAVODHDOI 480

Db 1097 KEESLTRASQYSQGLKVLIANTWLTSLKKEQDKTEILEAIESHHPRLASAVODHDOI 1156

Qy 481 VTSRKQEPAFHIAAGDAQLQKRMNVDSST 510

Db 1157 VTSRKQEPAFHIAAGDAQLQKRMNVDSST 1186

RESULT 7

ABJ37788

ID ABJ37788 standard; protein; 1349 AA.

XX AC ABJ37788;

XX XX

XX DT 15-MAY-2003 (first entry)

XX XX

XX DE Human tumour-related protein - SEQ ID No 573.

XX XX

XX KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;

XX KW tumour; breast cancer; cancer; immune response stimulation.

XX OS Homo sapiens.

XX XX

XX PN WO200283956-A1.

XX XX

XX PD 24-OCT-2002.

XX PF

XX PR 15-APR-2002; 2002WO-US012378.

XX PR

XX PR 13-APR-2001; 2001US-00834759.

XX PR 07-DEC-2001; 2001US-00007805.

XX PR 13-FEB-2002; 2002US-00076622.

XX XX

XX PA (CORI-) CORIXA CORP.

XX XX

XX PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;

XX PI Mitcham JR, Xu J, Harlocker SL, Hepler WI, Henderson RA, Fanger GR;

XX PI Vedvick TS, McNeill PD, Durham M;

XX XX

XX DR WPI: 2003-103376/09.

XX XX

XX PT New polypeptide and polynucleotide useful for stimulating and/or

XX PT expanding T cells specific for a tumor protein and treating breast

XX PT cancer.

XX PS

XX XX

XX XX Example 12; Page 353-357; 375pp; English.

XX CC

XX CC The invention comprises a method of stimulating and/or expanding T cells

XX CC specific for a tumour protein. The invention further comprises human

XX CC nucleic acids and proteins that are associated with tumours (e.g. breast

XX CC cancer). The method and sequences of the invention are useful for

XX CC stimulating and/or expanding T cells specific for a tumour protein,

XX CC detecting the presence of cancer, stimulating an immune response in a

XX CC patient and treating breast cancer. The present amino acid sequence

XX CC represents a human tumour-related protein

XX XX

XX SQ Sequence 1349 AA;

Query Match 99.6%; Score 2589; DB 6; Length 1349;

Best Local Similarity 100.0%; Pred. No. 3.4e-171; Indels 0; Gaps 0;

Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 60

Db 685 MKVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADIELPS 744

Qy 61 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120

Db 745 ESKQDYESSWDSSELCETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 804

Qy 121 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 180

Db 805 KVSIPTKALELMDMTFKAEPPKPSAFEPALIEQKSVNPKALELKNQETLRADQMPFSE 864

QY	181	SKQKVEENSWDSLSRETYSQDVCPKATHQKEMDKISGKLEDSTSLKILDTVHSCE	240
Db	865	SKQKVEENSWDSLSRETYSQDVCPKATHQKEMDKISGKLEDSTSLKILDTVHSCE	924
QY	241	RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	300
Db	925	RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN	984
QY	301	QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH	360
Db	985	QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH	1044
QY	361	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMTLKL	420
Db	1045	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMTLKL	1104
QY	421	KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIESHHPRLASAVQDHQI	480
Db	1105	KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIESHHPRLASAVQDHQI	1164
QY	481	VTSRKSQEPAFHIAGDACLQKKNVDVSST	510
Db	1165	VTSRKSQEPAFHIAGDACLQKKNVDVSST	1194
RESULT 8			
ID	ADL93235	standard; protein; 1349 AA.	
XX	ADL93235;		
XX	20-MAY-2004	(first entry)	
XX	Human breast cancer-associated polypeptide #53.		
XX	gene therapy; protein therapy; vaccine; breast cancer; cancer; human.		
XX	Homo sapiens.		
XX	US2003166022-A1.		
XX	04-SEP-2003.		
XX	15-APR-2002;	2002US-00124805.	
XX	28-DEC-1998;	98US-00222575.	
XX	02-APR-1999;	99US-00285480.	
XX	23-JUN-1999;	99US-00339338.	
XX	02-SEP-1999;	99US-00389681.	
XX	03-NOV-1999;	99US-00433826.	
XX	17-APR-2000;	2000US-005551621.	
XX	08-JUN-2000;	2000US-00590751.	
XX	22-JUN-2000;	2000US-00604287.	
XX	20-JUL-2000;	2000US-00620405.	
XX	13-APR-2001;	2001US-00834759.	
XX	07-DEC-2001;	2001US-00007805.	
XX	13-FEB-2002;	2002US-00076622.	
XX	(CORI-) CORIXA CORP.		
XX	Houghton RL, Sleath PR, Persing DH;		
XX	WPI; 2003-874918/81.		
XX	N-PSDB; ADL93228.		
XX	An isolated oncogenic polypeptide useful for preventing, diagnosing and treating breast cancer.		
XX	Example 12; SEQ ID NO 573; 294pp; English.		
XX	The invention relates to an isolated breast cancer-associated polypeptide. The polypeptide may be used for the diagnosis and treatment of breast cancers. The methods are useful for detecting the presence of a		

CC	cancer in a patient and treating a cancer in a patient. The present	
CC	sequence represents the amino acid sequence of a human breast cancer-	
CC	associated polypeptide.	
XX	Sequence 1349 AA;	
SQ	Query Match	99.6%; Score 2589; DB 7; Length 1349;
	Best Local Similarity	100.0%; Pred. No. 3.4e-171;
	Matches 510; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MKVSIPPTKALELMDMOTFKAEPPKPSAFEPALTEMQKSVPNKALELKNQETLRADQIFPS 60
Db	685	MKVSIPPTKALELMDMOTFKAEPPKPSAFEPALTEMQKSVPNKALELKNQETLRADQIFPS 744
QY	61	ESQKDYESSWDSLSRETYSQDVCPKATHQKIDKINGKLEESPDNDGFLKAPCRM 120
Db	745	ESQKDYESSWDSLSRETYSQDVCPKATHQKIDKINGKLEESPDNDGFLKAPCRM 804
QY	121	KVSIPPTKALELMDMOTFKAEPPKPSAFEPALTEMQKSVPNKALELKNQETLRADQIFPS 180
Db	805	KVSIPPTKALELMDMOTFKAEPPKPSAFEPALTEMQKSVPNKALELKNQETLRADQIFPS 864
QY	181	SKQKVEENSWDSLSRETYSQDVCPKATHQKIDKINGKLEESPDNDGFLKAPCRM 240
Db	865	SKQKVEENSWDSLSRETYSQDVCPKATHQKIDKINGKLEESPDNDGFLKAPCRM 924
QY	241	RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
Db	925	RARELQDHCEQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 984
QY	301	QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db	985	QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 1044
QY	361	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMTLKL 420
Db	1045	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNAELOMTLKL 1104
QY	421	KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIESHHPRLASAVQDHQI 480
Db	1105	KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIESHHPRLASAVQDHQI 1164
QY	481	VTSRKSQEPAFHIAGDACLQKKNVDVSST 510
Db	1165	VTSRKSQEPAFHIAGDACLQKKNVDVSST 1194
RESULT 9		
ID	AAU33351	standard; protein; 1002 AA.
XX	AAU33351;	
XX	18-DEC-2001	(first entry)
XX	Human breast cancer protein B726P from alternatively spliced cDNA.	
XX	Human; ss; breast cancer protein; tumour; cancer; cytostatic;	
XX	gene therapy; immunogen.	
XX	Homo sapiens.	
XX	WO200179286-A2.	
XX	25-OCT-2001.	
XX	12-APR-2001;	2001WO-US012164.
XX	17-APR-2000;	2000US-00551621.
XX	08-JUN-2000;	2000US-00590751.
XX	22-JUN-2000;	2000US-00604287.
XX	20-JUL-2000;	2000US-00620405.

PA (CORI-) CORIXA CORP.
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI: 2001-611721/70.
DR N-PSDB; AAS47411.
XX Breast Tumor Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer.
XX Claim 3; Page 281-283; 297pp; English.
XX The invention relates to isolated breast tumour proteins and nucleic
CC acids that encode them, including immunogenic fragments of the proteins.
CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
XX SQ Sequence 1002 AA;

Query Match 99.2%; Score 2578; DB 4; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTQTFKAEPPKPSAFAPAIEQMOKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTQTFKAEPPKPSAFAPAIEQMOKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQKDYESSWDSSELCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQKDYESSWDSSELCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTQTFKAEPPKPSAFAPAIEQMOKSVNPKALELKNQETLRADQMPFSE 180
Db 458 KVSIPTKALELMDMTQTFKAEPPKPSAFAPAIEQMOKSVNPKALELKNQETLRADQMPFSE 517
Qy 181 SKQKXVEENSWSSELRTVSQKDVCPKATHQKEMDKISGKLEBDSTSLKILDTVHSC 240
Db 518 SKQKXVEENSWSSELRTVSQKDVCPKATHQKEMDKISGKLEBDSTSLKILDTVHSC 577
Qy 241 RARELQKDHCHQRTGKQKBMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
Db 578 RARELQKDHCHQRTGKQKBMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 637
Qy 301 QEEKRRNADILNEKIREELGRIBEQHREKLEVKQQLQALRIQDIELKSVESNLNQVSH 360
Db 638 QEEKRRNADILNEKIREELGRIBEQHREKLEVKQQLQALRIQDIELKSVESNLNQVSH 697
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKLKEKNAELQMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKLKEKNAELQMTLKL 757
Qy 421 KEESLTKRASQVSGQLKVLIAENTMLTSLKKEKODKEILEAIESHHPRLASAVQDHQOI 480
Db 758 KEESLTKRASQVSGQLKVLIAENTMLTSLKKEKODKEILEAIESHHPRLASAVQDHQOI 817
Qy 481 VTSRKSQBPAPFHIAAGDAQLQRKMNVDVSST 510

Db 818 VTSRKSQBPAPFHIAAGDAQLQRKMNVDVSST 847
RESULT 10
ABG78918
ID ABG78918 standard; protein; 1002 AA.
XX AC ABG78918;
XX DT 15-NOV-2002 (first entry)
XX DE Human breast tumour polypeptide #10.
XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX OS Homo sapiens.
XX PN US2002085998-A1.
XX PD 04-JUL-2002.
XX PF 13-APR-2001; 2001US-00834759.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX FI Henderson RA;
XX DR WPI: 2002-635657/68.
XX N-PSDB; ABS64012.
XX PT Novel breast cancer polynucleotides and polypeptides encoded by the
PT polynucleotides, useful for detecting the presence of breast cancer in a
PT patient, and in pharmaceutical compositions, for treating breast cancer.
XX PS Claim 2; Page 208-211; 247pp; English.
XX CC The invention relates to an isolated breast tumour polynucleotide and the
CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC detecting the presence of breast cancer in a patient, and in
CC pharmaceutical compositions for treating breast cancer. The sequences are
CC useful for stimulating an immune response in a patient and can therefore
CC be used in production of vaccines. The sequences are also useful for
CC detecting the presence of a cancer in a patient, by obtaining a
CC biological sample from the patient, contacting the biological sample with
CC a composition of the invention and detecting the amount of polynucleotide
CC that hybridizes to the sample. This sequence represents a human breast
CC tumour polypeptide of the invention
XX SQ Sequence 1002 AA;

Query Match 99.2%; Score 2578; DB 5; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKVSIPTKALELMDMTQTFKAEPPKPSAFAPAIEQMOKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTQTFKAEPPKPSAFAPAIEQMOKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQKDYESSWDSSELCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQKDYESSWDSSELCTETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 457

Tue Feb 22 09:49:10 2005

us-09-602-362e-16.rag

Qy	121	KVSIPTKALELMDMOTFKAEPPKPSAFEPAPAEIEMQKSVNPKALELNEQTLRADQMPFSE	180
Db	458	KVSIPTKALELMDMOTFKAEPPKPSAFEPAPAEIEMQKSVNPKALELNEQTLRADQMPFSE	517
Qy	181	SKQKXVEENSWDSLSLRETQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE	240
Db	518	SKQKXVEENSWDSLSLRETQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE	577
Qy	241	RARELOKDHCEORTGKMEQKKFCVLKKLSEAKEIKSOLENKVKEQELCSVRLTLN	300
Db	578	RARELOKDHCEORTGKMEQKKFCVLKKLSEAKEIKSOLENKVKEQELCSVRLTLN	637
Qy	301	QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNVSH	360
Db	638	QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNVSH	697
Qy	361	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL	420
Db	698	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL	757
Qy	421	KEESLTRKASQVSGQLKVLIAENTMLTSLKLEKQDKEILEAEIESHHPRLASAVQDHDQI	480
Db	758	KEESLTRKASQVSGQLKVLIAENTMLTSLKLEKQDKEILEAEIESHHPRLASAVQDHDQI	817
Qy	481	VTSRKSQBPAPHIAGDACLQRKNVDVSST	510
Db	818	VTSRKSQBPAPHIAGDACLQRKNVDVSST	847
RESULT 11			
ABJ37741	ID	ABJ37741 standard; protein; 1002 AA.	
XX	AC	ABJ37741;	
XX	DT	15-MAY-2003 (first entry)	
XX	DE	Human tumour-related protein - SEQ ID No 475.	
XX	KW	Human; vaccine; gene therapy; T cell stimulation; T cell expansion;	
XX	KW	tumour; breast cancer; cancer; immune response stimulation.	
XX	OS	Homo sapiens.	
XX	PN	WO200283956-A1.	
XX	PD	24-OCT-2002.	
XX	PF	15-APR-2002; 2002WO-US012378.	
XX	PR	13-APR-2001; 2001US-00834759.	
XX	PR	07-DEC-2001; 2001US-00007805.	
XX	PR	13-FEB-2002; 2002US-00076622.	
XX	PA	(CORI-) CORIXA CORP.	
XX	PI	Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;	
XX	PI	Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;	
XX	PI	Vedvick TS, McNeill PD, Durham M;	
XX	XX	WPI; 2003-103376/09.	
XX	XX	New polypeptide and polynucleotide useful for stimulating and/or	
XX	PT	expanding T cells specific for a tumor protein and treating breast	
XX	PT	cancer.	
XX	PS	Example 1; Page 298-300; 375pp; English.	
XX	XX	The invention comprises a method of stimulating and/or expanding T cells	
XX	XX	specific for a tumour protein. The invention further comprises human	
XX	XX	nucleic acids and proteins that are associated with tumours (e.g. breast	
XX	XX	cancer). The method and sequences of the invention are useful for	
stimulating and/or expanding T cells specific for a tumour protein, a			
detecting the presence of cancer, stimulating an immune response in a			
patient and treating breast cancer. The present amino acid sequence			
represents a human tumour-related protein			
XX	XX	Sequence 1002 AA;	
Qy	1	MKVSIPPTKALELMDMOTFKAEPPKPSAFEPAPAEIEMQKSVNPKALELNEQTLRADQMPFSE	60
Db	338	MKVSIPPTKALELMDMOTFKAEPPKPSAFEPAPAEIEMQKSVNPKALELNEQTLRADQMPFSE	397
Qy	61	ESKQDYESSWDSLSLRETQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE	120
Db	398	ESKQDYESSWDSLSLRETQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE	457
Qy	121	KVSIPTKALELMDMOTFKAEPPKPSAFEPAPAEIEMQKSVNPKALELNEQTLRADQMPFSE	180
Db	458	KVSIPTKALELMDMOTFKAEPPKPSAFEPAPAEIEMQKSVNPKALELNEQTLRADQMPFSE	517
Qy	181	SKQKXVEENSWDSLSLRETQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE	240
Db	518	SKQKXVEENSWDSLSLRETQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE	577
Qy	241	RARELOKDHCEORTGKMEQKKFCVLKKLSEAKEIKSOLENKVKEQELCSVRLTLN	300
Db	578	RARELOKDHCEORTGKMEQKKFCVLKKLSEAKEIKSOLENKVKEQELCSVRLTLN	637
Qy	301	QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNVSH	360
Db	638	QEEKRRNADILNEKIREELGRIBEOHRKELEVKQLEQALRIQDIELKSVESNLNVSH	697
Qy	361	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL	420
Db	698	THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKL	757
Qy	421	KEESLTRKASQVSGQLKVLIAENTMLTSLKLEKQDKEILEAEIESHHPRLASAVQDHDQI	480
Db	758	KEESLTRKASQVSGQLKVLIAENTMLTSLKLEKQDKEILEAEIESHHPRLASAVQDHDQI	817
Qy	481	VTSRKSQBPAPHIAGDACLQRKNVDVSST	510
Db	818	VTSRKSQBPAPHIAGDACLQRKNVDVSST	847
RESULT 12			
ADL93137	ID	ADL93137 standard; protein; 1002 AA.	
XX	AC	ADL93137;	
XX	DT	20-MAY-2004 (first entry)	
XX	DE	Human breast cancer-associated polypeptide #10.	
XX	KW	gene therapy; protein therapy; vaccine; breast cancer; cancer; human.	
XX	OS	Homo sapiens.	
XX	PN	US2003166022-A1.	
XX	PD	04-SEP-2003.	
XX	PF	15-APR-2002; 2002US-00124805.	
XX	PR	28-DEC-1998; 98US-00222575.	
XX	PR	02-APR-1999; 99US-00285480.	
XX	PR	23-JUN-1999; 99US-00339338.	
XX	PR	02-SEP-1999; 99US-00389681.	
XX	PR	03-NOV-1999; 99US-00433826.	

PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH;
XX WPI; 2003-874918/81.
XX N-PSDB; ADL93136.
XX An isolated oncogenic polypeptide useful for preventing, diagnosing and
XX treating breast cancer.
XX Example 1; SEQ ID NO 475; 294pp; English.
XX The invention relates to an isolated breast cancer-associated
XX polypeptide. The polypeptide may be used for the diagnosis and treatment
XX of breast cancers. The methods are useful for detecting the presence of a
XX cancer in a patient and treating a cancer in a patient. The present
XX sequence represents the amino acid sequence of a human breast cancer-
XX associated polypeptide.
XX Sequence 1002 AA;

Query Match 99.2%; Score 2578; DB 7; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKAXHQEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADMFPS 180
Db 458 KVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADMFPS 517
Qy 181 SKQKXVENSWSLSRETQSKDVCPKATHQKEMDKISGLSDSTLSKILDTVHSC 240
Db 518 SKQKXVENSWSLSRETQSKDVCPKATHQKEMDKISGLSDSTLSKILDTVHSC 577
Qy 241 RARELQKDHCHQRTGRMEQMKKFCVLKKKLSEAKEIKSQLENQKVKWEQELCSVRLTN 300
Db 578 RARELQKDHCHQRTGRMEQMKKFCVLKKKLSEAKEIKSQLENQKVKWEQELCSVRLTN 637
Qy 301 QEEKRRNADILNKIREELGRIBEOHRKELEVQKQLEQALRIQDIELKSVEHNLQVSH 360
Db 638 QEEKRRNADILNKIREELGRIBEOHRKELEVQKQLEQALRIQDIELKSVEHNLQVSH 697
Qy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNKAELOMTLKL 420
Db 698 THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKKNKAELOMTLKL 757
Qy 421 KEESLTKRASQYSQLKVLIANTMTLSKLKEKQDKILEAEIESHPRLASAVQDDHQI 480
Db 758 KEESLTKRASQYSQLKVLIANTMTLSKLKEKQDKILEAEIESHPRLASAVQDDHQI 817
Qy 481 VTSRKSQEPAPHIAGDAQLQRKNVDVSS 510
Db 818 VTSRKSQEPAPHIAGDAQLQRKNVDVSS 847
RESULT 13
ADE44427
ID ADE44427 standard; protein; 1002 AA.

XX ADE44427;
XX 29-JAN-2004 (first entry)
XX Human breast cancer protein #10.
XX human; breast tumour; cancer; vaccine; T cell stimulator;
XX T cell expander.
XX Homo sapiens.
XX US2003104366-A1.
XX 05-JUN-2003.
XX 17-APR-2000; 2000US-00551621.
XX 28-DEC-1998; 98US-00222575.
XX 02-APR-1999; 99US-00285480.
XX 23-JUN-1999; 99US-00339338.
XX 02-SEP-1999; 99US-00389681.
XX 03-NOV-1999; 99US-00433826.
XX (JIAN/) JIANG Y.
XX (DILL/) DILLON D C.
XX (MITC/) MITCHAM J L.
XX (XUJJ/) XU J.
XX (HARL/) HARLOCKER S L.
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX WPI; 2004-020270/02.
XX N-PSDB; ADE44426.
XX Novel isolated polypeptide comprising immunogenic portion of breast tumor
XX protein or its variant, useful for formulating vaccines for inhibiting
XX cancer development in a patient.
XX Example 1; SEQ ID NO 475; 217pp; English.
XX The invention relates to an isolated polypeptide comprising at least an
XX immunogenic portion of a breast tumour protein. The polynucleotide, its
XX polypeptide, its antibody, a pharmaceutical composition comprising the
XX fusion protein or the polynucleotide encoding it, a vaccine comprising
XX the fusion protein or the polynucleotide encoding it, an isolated T cell
XX population comprising T cells specific for a breast tumour protein, and a
XX method for removing tumour cells from a biological sample is useful for
XX inhibiting the development of a cancer in a patient. The polypeptide is
XX useful for stimulating and/or expanding T cells specific for a breast
XX tumour protein. Stimulating and/or expanding T cells specific for a
XX breast tumour protein is useful for inhibiting the development of a
XX cancer in a patient. The method additionally involves the step of cloning
XX at least one proliferated cell and then administering the cloned T cells
XX to the patient. The present sequence represents a breast cancer protein.
XX Sequence 1002 AA;

Query Match 99.2%; Score 2578; DB 8; Length 1002;
Best Local Similarity 99.6%; Pred. No. 1.3e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADIELPS 60
Db 338 MKVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADIELPS 397
Qy 61 ESKQDYEESSWDSLSLCTVTSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 398 ESKQDYEESSWDSLSLCTVTSQKDVCLPKAXHQEIDKINGKLEESPDNDGFLKAPCRM 457
Qy 121 KVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADMFPS 180
Db 458 KVSIPTKALELMDMTFKAPPPKPSAFPAIEWQKSVNPKALELKNQETLRADMFPS 517

CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
CC
SQ Sequence 1095 AA;
Query Match 99.2%; Score 2578; DB 4; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.5e-170; Indels 0; Gaps 0;
Matches 508; Conservative
QY 1 MKVSIPTKALELMDMOTFKAEPPKPSAFEPKPAIEEMQKSVNKALELKNQTLRADPILPS 60
Db 431 MKVSIPTKALELMDMOTFKAEPPKPSAFEPKPAIEEMQKSVNKALELKNQTLRADPILPS 490
QY 61 ESKQDYESSWDSSESLCETVSKOVCLPKATHQKIDKINGKLESPNDGFLKAPCRM 120
Db 491 ESKQDYESSWDSSESLCETVSKOVCLPKATHQKIDKINGKLESPNDGFLKAPCRM 550
QY 121 KVSIPPTKALELMDMOTFKAEPPKPSAFEPKPAIEEMQKSVNKALELKNQTLRADPILPS 180
Db 551 KVSIPPTKALELMDMOTFKAEPPKPSAFEPKPAIEEMQKSVNKALELKNQTLRADPILPS 610
QY 181 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 240
Db 611 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 670
QY 241 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 300
Db 671 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 730
QY 301 QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
Db 731 QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 790
QY 361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
Db 791 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 850
QY 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
Db 851 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 910
QY 481 VTSRKSQEPAFHAGDACLQKMNVDVSST 510
Db 911 VTSRKSQEPAFHAGDACLQKMNVDVSST 940
RESULT 15
ABG78924
ID ABG78924 standard; protein; 1095 AA.
XX AC ABG78924;
XX DT 15-NOV-2002 (first entry)
XX DE Human breast tumour polypeptide #15.
XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX OS Homo sapiens.
XX PN US2002085998-A1.
XX PP 04-JUL-2002.
CC

181 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 240
518 SKQKVEENSWSSESLRETYSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCE 577
241 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 300
578 RARELQKHCEORTGKMEQMKKFCVLKKLSEAKETKSOLENQKVKWEQELCSVRLTLN 637
301 QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 360
638 QEEKRRNADILNEKIREELGRIEEQHKELEVKQLEQALRIQDIELKSVESNLNQVSH 697
361 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 420
698 THENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKL 757
421 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 480
758 KEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKKEILEAEIESHHPRLASAVQDHDQI 817
481 VTSRKSQEPAFHAGDACLQKMNVDVSST 510
818 VTSRKSQEPAFHAGDACLQKMNVDVSST 847
RESULT 14
AAU33357
ID AAU33357 standard; protein; 1095 AA.
XX AC AAU33357;
XX DT 18-DEC-2001 (first entry)
XX DE Human breast cancer protein B726P fusion protein #1.
XX KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX OS Homo sapiens.
XX PN WO200179286-A2.
XX PD 25-OCT-2001.
XX PF 12-APR-2001; 2001WO-US012164.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PA (CORI-) CORIXA CORP.
XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
XX DR N-PSDB; AAS47421.
XX PT Breast Tumor Proteins and nucleic acids useful for the prevention,
XX diagnosis and treatment of breast cancer.
XX PS Claim 22; Page 292-295; 297pp; English.
XX CC The invention relates to isolated breast tumour proteins and nucleic
CC acids that encode them, including immunogenic fragments of the proteins.
CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g. by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect

Oy 421 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHQOI 480
Db 851 KEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKKEILEAEIESHHPRLASAVQDHQOI 910
Oy 481 VTSRKSQEPAPHIAGDACLQKRMVDSST 510
Db 911 VTSRKSQEPAPHIAGDACLQKRMVDSST 940

Search completed: February 19, 2005, 02:04:38
Job time : 58.7765 secs

XX 13-APR-2001; 2001US-00834759.
XX 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX (CORI-) CORIXA CORP.
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI Henderson RA;
XX WPI; 2002-635657/68.
DR N-PSDB; ABS64022.
XX Novel breast cancer polynucleotides and polypeptides encoded by the
PT polynucleotides, useful for detecting the presence of breast cancer in a
PT patient, and in pharmaceutical compositions, for treating breast cancer.
XX Disclosure; Page 223-225; 247pp; English.
XX The invention relates to an isolated breast tumour polynucleotide and the
CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC detecting the presence of breast cancer in a patient, and in
CC pharmaceutical compositions for treating breast cancer. The sequences are
CC useful for stimulating an immune response in a patient and can therefore
CC be used in production of vaccines. The sequences are also useful for
CC detecting the presence of a cancer in a patient, by obtaining a
CC biological sample from the patient, contacting the biological sample with
CC a composition of the invention and detecting the amount of polynucleotide
CC that hybridizes to the sample. This sequence represents a human breast
XX tumour polypeptide of the invention
XX
SQ Sequence 1095 AA;
Query Match 99.2%; Score 2578; DB 5; Length 1095;
Best Local Similarity 99.6%; Pred. No. 1.5e-170;
Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 MKVSIPTKALELMDMTFKAPPEKPSAFEPALIEQKSVNPKALELKNQTLRADEILPS 60
Db 431 MKVSIPTKALELMDMTFKAPPEKPSAFEPALIEQKSVNPKALELKNQTLRADEILPS 490
Oy 61 ESKQDYESSWDSSESLCETYSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 120
Db 491 ESKQDYESSWDSSESLCETYSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRM 550
Oy 121 KVSIPTKALELMDMTFKAPPEKPSAFEPALIEQKSVNPKALELKNQTLRADMPFSE 180
Db 551 KVSIPTKALELMDMTFKAPPEKPSAFEPALIEQKSVNPKALELKNQTLRADMPFSE 610
Oy 181 SKQKXVENSWSSESLRETTSQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE 240
Db 611 SKQKXVENSWSSESLRETTSQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCE 670
Oy 241 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 300
Db 671 RARELQKDHCQRTGKMEQMKKFCVLKKLSEAKEIKSQLENQKVKWEQELCSVRLTLN 730
Oy 301 QEEKRRNADILNEKIRELGRIBEQHRLKEVQQLQALRIQDIELKSVESNLNQVSH 360
Db 731 QEEKRRNADILNEKIRELGRIBEQHRLKEVQQLQALRIQDIELKSVESNLNQVSH 790
Oy 361 THENENYLLHNCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 420
Db 791 THENENYLLHNCMLKKEIAMLKLEIATLKHOYEKENKYPEDIKILKEKNAELQMTLKL 850

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:55:34 ; Search time 124.257 Seconds
(without alignments)
4166.462 Million cell updates/sec

Title: US-09-602-362E-27
Perfect score: 5173
Sequence: 1 MVATLLSYGAVIEYQNKASL.....REQLKVLTAENTMLTSLKE 1011

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_eprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5173	100.0	1011	2 Q9BXX2	Q9bxx2 homo sapien
2	3075	59.4	1341	2 Q9BXX3	Q9bxx3 homo sapien
3	1011.5	19.6	424	2 Q9NS19	Q9ns19 homo sapien
4	923	17.8	1715	2 Q9UPS8	Q9ups8 homo sapien
5	920.5	17.8	1710	2 Q9HIQ1	Q9hiq1 homo sapien
6	703	13.6	1080	2 Q9ZRI4	Q9zri4 homo sapien
7	553	10.7	168	2 Q9N7B8	Q9n7b8 homo sapien
8	489	9.5	718	2 Q9HCD1	Q9hcd1 homo sapien
9	465	9.0	1043	2 Q9ZS22	Q9zs22 mus musculu
10	396	7.7	823	2 Q9H0H6	Q9h0h6 homo sapien
11	359	6.9	2006	2 Q7K5Q6	Q7k5q6 plasmodium
12	359	6.9	2019	2 Q7K5Q5	Q7k5q5 plasmodium
13	359	6.9	2055	2 Q8IHP3	Q8ihp3 plasmodium
14	359	6.9	2055	2 Q8TSC7	Q8tsc7 plasmodium
15	357.5	6.9	10578	2 Q8ISF5	Q8isf5 caenorhabdi
16	357.5	6.9	18519	2 Q8ISP6	Q8isf6 caenorhabdi
17	357.5	6.9	18534	2 Q8ISP7	Q8isf7 caenorhabdi
18	346.5	6.7	2042	2 Q8BUQ9	Q8buq9 debaryomyce
19	346	6.7	1416	2 Q9BZF9	Q9bzf9 homo sapien
20	346	6.7	1449	2 Q9BG87	Q9bg87 bos taurus
21	345	6.7	1416	2 Q9HCL1	Q9hcl1 homo sapien
22	342.5	6.6	1980	2 Q6FWE0	Q6fwe0 candida gla
23	341.5	6.6	1427	1 R8ST_HUMAN	P30622 homo sapien
24	340.5	6.6	1268	2 Q7LGT1	Q7lgt1 saccharomyc
25	340.5	6.6	1790	2 Q07380	Q07380 saccharomyc
26	340	6.6	1413	2 Q8CGB3	Q8cgb3 mus musculu
27	338.5	6.5	1790	1 U501_YEAST	P25386 saccharomyc
28	337	6.5	641	2 Q60311	Q60311 homo sapien
29	336	6.5	453	2 Q8IZM7	Q8izm7 homo sapien
30	336	6.5	992	1 AN18_HUMAN	Q8ivf6 homo sapien
31	331.5	6.4	7210	2 Q9V7G8	Q9v7g8 drosophila

ALIGNMENTS

RESULT 1

ID	Q9BXX2	PRELIMINARY;	PRT;	1011 AA.
AC	Q9BXX2			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-JUN-2003 (Tremblrel. 24, Last annotation update)			
DE	Breast cancer antigen NY-BR-1.1 (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=21174979; PubMed=11280766;			
RA	Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,			
RA	Jaeger E., Knuth A., Old L.J., Chen Y.T.;			
RT	"Identification of a Tissue-specific Putative Transcription Factor in			
RT	Breast Tissue by Serological Screening of a Breast Cancer Library."			
RL	Cancer Res. 61:2055-2061(2001)			
DR	EMBL; AF269088; AAK27326.1;			
DR	HSSP; O75832; IUOH.			
DR	InterPro; IPR002110; ANK.			
DR	Pfam; PF00023; Ank; 3.			
DR	SMART; SM00248; ANK; 3.			
DR	PROSITE; PS50088; ANK_REPEAT; 1.			
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.			
KW	ANK repeat.			
FT	NON TER 1011 1011			
SQ	SEQUENCE 1011 AA; 114250 MW; 16627D7B218D8438 CRC64;			
Query Match 100.0%; Score 5173; DB 2; Length 1011;				
Best Local Similarity 100.0%; Pred. No. 1.6e-194;				
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	MVATLLSYGAVIEYQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESCKTALMLAI	60	
Db	1	MVATLLSYGAVIEYQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESCKTALMLAI	60	
Qy	61	CEGSSEIVGMLLQNVDFVPAEDIHGITAERYAAAGVNYIHQQLLEHIRKLPKIPQNTNP	120	
Db	61	CEGSSEIVGMLLQNVDFVPAEDIHGITAERYAAAGVNYIHQQLLEHIRKLPKIPQNTNP	120	
Qy	121	EGTSTGTPDEAAPLAERTPDTABSLLEKTPDEAARLVEGTSKAKIQCLGKATSGFEQSTE	180	
Db	121	EGTSTGTPDEAAPLAERTPDTABSLLEKTPDEAARLVEGTSKAKIQCLGKATSGFEQSTE	180	
Qy	181	ETPRKILRPKTSEKFSWPAKERSRKITWEEKTSVKTECVAGVTPNKTEVLEKGTSM	240	
Db	181	ETPRKILRPKTSEKFSWPAKERSRKITWEEKTSVKTECVAGVTPNKTEVLEKGTSM	240	
Qy	241	IACPTKETSTKASTNDVSSVEPIFSLFGTRTINSQCTKVEEDFNLATKIISKSAQNY	300	
Db	241	IACPTKETSTKASTNDVSSVEPIFSLFGTRTINSQCTKVEEDFNLATKIISKSAQNY	300	

QY 301 TCLPDATYQKDIKTINHIEDQMPSPSKREDEDEYSDWDSGLSPSSAKTOVCIPESMYQ 360
 DB 301 TCLPDATYQKDIKTINHIEDQMPSPSKREDEDEYSDWDSGLSPSSAKTOVCIPESMYQ 360
 QY 361 KVMENREVELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPSPSKQKDEEN 420
 DB 361 KVMENREVELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPSPSKQKDEEN 420
 QY 421 SWDSSEPCETVSKQDVYLPKATHQKEFTLSSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
 DB 421 SWDSSEPCETVSKQDVYLPKATHQKEFTLSSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
 QY 481 LKDRFTFAESPDKDGLLKPTCGRKVSLPNKALEKURETILKASPDNDGLLKPTCGRKV 540
 DB 481 LKDRFTFAESPDKDGLLKPTCGRKVSLPNKALEKURETILKASPDNDGLLKPTCGRKV 540
 QY 541 SLPNKALELKDRFTFAESPDKDGLLKPTCGRKVSLPNKALEKURETILKASPDNDGLLKPTCGRKV 600
 DB 541 SLPNKALELKDRFTFAESPDKDGLLKPTCGRKVSLPNKALEKURETILKASPDNDGLLKPTCGRKV 600
 QY 601 DTLSGKLEESPDKDGLLKPTCGRKVSLPNKALELKURETILKASPDNDGLLKPTCGRKV 660
 DB 601 DTLSGKLEESPDKDGLLKPTCGRKVSLPNKALELKURETILKASPDNDGLLKPTCGRKV 660
 QY 661 ENSQTKVEEDPNLTTKGATKTVTGQQRDGIIGIERAPOQTNKMPSELGRKEDTKST 720
 DB 661 ENSQTKVEEDPNLTTKGATKTVTGQQRDGIIGIERAPOQTNKMPSELGRKEDTKST 720
 QY 721 SDSEIISVSDTONYECLEPEATYQKEIKTTNGKIEESPEKSHFEPATEMNSVPNKGLEW 780
 DB 721 SDSEIISVSDTONYECLEPEATYQKEIKTTNGKIEESPEKSHFEPATEMNSVPNKGLEW 780
 QY 781 KVKQTLRADSTLTKILDALPSCERBELKKNCEQITAKMEQMKNCVILQKELSEAKE 840
 DB 781 KVKQTLRADSTLTKILDALPSCERBELKKNCEQITAKMEQMKNCVILQKELSEAKE 840
 QY 841 IKSQLENOQAKWEQELCSVRLPLNOBEEKRRNVDILKEKIRPEOLRKKLEVKHOLEQTL 900
 DB 841 IKSQLENOQAKWEQELCSVRLPLNOBEEKRRNVDILKEKIRPEOLRKKLEVKHOLEQTL 900
 QY 901 RIQDIELKSVTSNLNQVSHSENDLFHENCMLKEIAMLKLEVALTLKHQHVQKENVYF 960
 DB 901 RIQDIELKSVTSNLNQVSHSENDLFHENCMLKEIAMLKLEVALTLKHQHVQKENVYF 960
 QY 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011
 DB 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011

RESULT 2

Q9BXX3 PRELIMINARY; PRT; 1341 AA.
 ID Q9BXX3
 AC Q9BXX3;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Breast cancer antigen NY-BR-1.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21174979; PubMed=11280766;
 RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
 RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
 RT "Identification of a Tissue-specific Putative Transcription Factor in
 RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
 RL Cancer Res. 61:2035-2061 (2001).
 DR ENBL; AF269087; AAK27325.1; --
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0005515; F:protein binding; NAS.

DR GO; GO:0003700; F:transcription factor activity; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 6.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 6.
 DR PROSITE; PS00088; ANK_REPEAT; 4.
 DR PROSITE; PS02927; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
 KW ANK repeat.
 SQ SEQUENCE 1341 AA; 152776 MW; 33E53DD6FD3A58B CRC64;

Query Match 59.4%; Score 3075; DB 2; Length 1341;
 Best Local Similarity 57.4%; Pred. No. 2.3e-112;
 Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;
 QY 1 MVALTSLYGAIVIVQNKASLTPLLLAIQKRSKQTVFLLTKNANANAFVSKCTALMLAI 60
 DB 98 VVAKLLSHGAVIEVHNKASLTPLLLSITTKRSEQIVFLLTKNANANAVNKYKCTALMLAV 157
 QY 61 CEGSSEITVGMLOOVNDVFAEDIHGITAERYAAARGVNYIHQOLLEHRIKLPKNPONTNP 120
 DB 158 CHGSSSEITVGMLOOVNDVFAADICGVTAHYAVTCGFHHIHEQIMYIRKLSKHQNTNP 217
 QY 121 EGTSTGTDPDEAALA-----ERTPTDAESLLEKTPDEAARLVE 158
 DB 218 EGTSGAGTDEAALAERTPTDAESLVEKTPDEAALPVERPTDAESLVEKTPDEAASLVE 277
 QY 159 GTSAKIQCLOKATSGKPEQSTBTETPKILAPTTETSEKFSWP----- 200
 DB 278 GTSDKIQCLEKATSGKPEQSAEETPREITSFAKETSSEKFTWPAKGRPKTAWKEDTTPR 337
 QY 201 -----AKERSKITWEEKETSVKTECVAGTVPNKTEVLEKGTNNMIACP 244
 DB 338 EIMSPAKETSEKFTWAAKGRPKLAWKKEKTPVKTGCVARVTSNKTIVLEKGRSKMIACP 397
 QY 245 TKETSTKASTNDVSSVEPIPSLFGTRTNIENSQCTKVEEDFNLATKIISAKAANYTCLP 304
 DB 398 TKESSTKASAN----- 408
 QY 305 DATYQKDIKTINHIEDQMPSPSKREDEDEYSDWDSGLSPSSAKTOVCIPESMYQKWE 364
 DB 409 -----DQRFPSKQBEDEYSDWDSGLSPSSAKTOVCIPESMYQKWE 452
 QY 365 INREVEELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPSPSKQKDEENSWDS 424
 DB 453 INREVEELPEKPSAFKPAVEMQKTPVNPNAFELKNEQTLRAAQMPSPSKQKDEENSWDS 512
 QY 425 EPCETVSKQDVYLPKATHQKEFTLSSGKLEESPVKDGLLKPTCGRKVSLPNKALELKDR 484
 DB 513 ESLCETVSKQDVYLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMD 572
 QY 485 ETFKAEPSDKDGLLKPTCGRKVSLPNKALELKURETILKASPDNDGLLKPTCGRKVSLPN 544
 DB 573 QTFKAEPPGKPSAFEP-----ATENQK-----SVPN 598
 QY 545 KALELKDRFTFAKAAQMPSPSKQKDEENSWDSFLETLLQNDVCLPKATHQKEFTLS 604
 DB 599 KALELKNEQTLWRADEILLPESKQKDEENSWDSFLETLLQNDVCLPKATHQKEIDKIN 658
 QY 605 KGLBESPDKDGLLKPTCGMKISLPNKALELKDRFTFAEDVSSVSESTFSLFGKPTTNSQ 664
 DB 659 KGLBESPDKDGLLKATCGMKVSIPTKALELMDMTQFKABPPEK-PSAF-----EPAIENQK 713
 QY 665 STKVEEDPNLTTKGATKTVTGQQRDGIIGIERAPOQTNKMPSELGRKEDTKSTDS 724
 DB 714 SV-----PNKALELKNEQTLWRADEILLP-----SEKQKQDEESSWDS 751
 QY 725 IISVSDTONYECLEPEATYQKEIKTTNGKIEES----- 756
 DB 752 SLCEVTSQKDVCLPKATHQKEIDKINGKLEESPNKDGFLKAPCRMKVSIPTKALELMDMQ 811

```
Qy 757 -----PEKSPHFPATQMNSVFNKGLGWKQKOTLRA----- 789
Db 812 TFKAEPPEKPSAFPAIEMQKSVFNKALELNEQTLRQADQMPPSESQKQVVEENSWDS 871
Qy 789 -----DSTTLKSLDALPSCERGLKNDKCEQITA 819
Db 872 LRETVSQKQVCPKATQKQKEMDKISGLKEDSTLSKILDTVHSCERARELQKHCQRTG 931
Qy 820 KMEQKMKFCVLOKELSEAKEIKSOLENQKAKWQELCSVRLPLNQBEERKRNVDILKEK 879
Db 932 KMEQKMKFCVLLKKLSEAKEIKSOLENQKVKWQELCSVRLTLNQBEERKERNADILNEK 991
Qy 880 IRP-----EEQLRKLEVKHOLEQTLRIQDIQELKSVTSNLSNTHSENDLPHENCM 934
Db 992 IREBLGRIEEQRKELEVKQLEQALRIQDIQELKSVESNLSNTHSENNYLLHNCML 1051
Qy 935 KKETAMKLEVATLKHOHVKENYKFEDIKLOKNAELQMTLKKOKVTVKRASQYREQ 994
Db 1052 KKEIAMKLEIATLKHQIQEKENYKFEDIKILKRNALQMTLKKESLTKRASQYSGQ 1111
Qy 995 LKVLTAENTMLTSKLKE 1011
Db 1112 LKVLIAENTMLTSKLKE 1128

RESULT 3
Q9NS19 PRELIMINARY; PRT; 424 AA.
AC Q9NS19;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE PRED4 protein (fragment).
GN Name=PRED4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,
RA Toyoda A., Ishii K., Ito K., Choi D.K., Souda E., Ohki M.,
RA Takagi T., Sakai Y., Tauden S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Riessmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehmach H., Reinhardt R., Yaspo M.L.;
RU Submitted (MAY-2000) to the ENBL/GenBank/DBJ databases.
DR EMBL; AL163203; CAB9394.1; -.
FT NON TER 1
FT NON TER 424
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

Query Match 19.6%; Score 1011.5; DB 2; Length 424;
Best Local Similarity 91.5%; Pred. No. 2.9e-32;
Matches 204; Conservative 6; Mismatches 34; Indels 9; Gaps 1;

Qy 789 DSTTSLKILDALPSCERGLKNDKCEQITAKMQKNKFCVLOKELSEAKEIKSOLENQ 848
Db 2 DSTTSLKILDALPSCERGLKNDKCEQITAKMQKNKFCVLOKELSEAKEIKSOLENQ 61
Qy 849 KAKWEQELCSVRLPLNQBEERKRNVDILKEKIRPEEQKLEVKHOLEQTLRIQDIQELK 908
Db 62 KAKWEQELCS-----KKRRDVIDILKEKIRPEEQKLEVKHOLEQALRIQDIQELK 112
Qy 909 SVTSNLSNTHSENDLPHENCMKKEIAMKLEIATLKHQVQKENVKPFEDIKILQ 969
Db 113 SVTSNLSNTHSENDLPHENCMKKEIAMKLEIATLKHQVQKENVKPFEDIKILQ 172
```

```
Qy 969 KNAELQMTLKLKOKVTVKRASQYREQKLVLTAEINTMLTSKLKE 1011
Db 173 KNAELQMTLKLKOKVTVKRASQYREQKLVLTAEINTMLTSKLKE 215

RESULT 4
Q9UPS8 PRELIMINARY; PRT; 1715 AA.
AC Q9UPS8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE KIAA1074 protein (fragment).
GN Name=KIAA1074;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RU DNA Res. 6:197-205(1999).
DR EMBL; AB028997; BAA83026.2; -.
DR HSP; P20749; IK1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS00088; ANK_REPEAT; 4.
DR PROSITE; PS00297; ANK_REPEAT; 1.
FT ANK repeat.
FT NON TER 1
FT NON TER 1715
SQ SEQUENCE 1715 AA; 196962 MW; A82858F5F58E3203 CRC64;

Query Match 17.8%; Score 923; DB 2; Length 1715;
Best Local Similarity 28.5%; Pred. No. 4.2e-28;
Matches 320; Conservative 157; Mismatches 342; Indels 304; Gaps 34;

Qy 7 SYGAVIEVQ-----NKASLTPLLAIQKRSKOTVEPLLTKNANANAFNSKC 53
Db 61 SAGNVAKVQQLLLRNKGLNDROKNTALHLACANGHPEVVTLLVDKCOLNVCDNENR 120
Qy 54 TALMLAICEGSSIEVGMQLQONVDVFAEDIHGITAERYAAARGVNYIHQQLLEHRLPK 113
Db 121 TALMKAVQCCEKCATILSHGADPNLADVHGNTALHYA-----VYNEISVATKLL 173
Qy 114 NPQNTNPGTSTGTPDEAAPIAERTPTAESLLEKTDEARLVEGTSIAKICLGKATSG 173
Db 174 YDANIEAK-----NKODLTPL-----LL-----AVSG 195
Qy 174 KFGSTEETPRKILRPKTESKESWPAKRSKRITWEEKETSVKTECVAGVTNPKTEVL 233
Db 196 KQOQWVEFLIK--KANVNAVDKL-----ESSHOLISEYKEER----- 232
Qy 234 EKGTSNMIACTKETSTKASTNVDSVVEPISFLGTRTIENTSQCTKVEEDFNLATKIIS 293
Db 233 -----PKHSSNSNVDESSEDSLSRLSGKPGVDDSWPTDDEDLNFTDNVP 280
Qy 294 K-----SAAQNYTCLPDATYQDKIKTINHKI--EDMFPSESKEDEEYSWDSGLPE 345
Db 281 KPSLAKLMTASQOSRKNLEATY-GTVRTGNRTLPED-----RSDSDQDEVVVSLEPT 331
Qy 346 SSAKTQVCIPESMYQ-----KVMENRREVEELPEKPSAFKP--AVEMQKTVNKA 393
Db 332 TSIRKQV-CFSPHTYQSPDLLPKPSHKSHPANGLMKKEPTKPGIAKNGIDILIESAP--- 387
```

```
QY 394 FELKNEOTL-----RAAOMFPSESOKODDENSWSESPCTVSKQDVLPKATH 443
Db 388 LEQTNNDNLTVYDVHKNRDMMSALGLGQEEIDESPWSESIENFPQKYV----- 440
QY 444 QKEFDTLGKLESPVKGDLKPTCGRKVSLPNKALELKORETPKASSPKDGLLKTGCG 503
Db 441 ----DPLAGAAD-----GKEKNIGNEQAE-----DVFYIPSC- 468
QY 504 RKVSLPNKALELKORETL-----KAESPNDGLKPTCGRKVSLPNKALELKORETPKAA 558
Db 469 MSGSRNFMAKLEDTNRVGMVAHMESEPERYLHLKPTIEMKDSVPNKAGGKMDVQTSKAA 528
QY 559 Q-----MFPSESOKODDENSWDFESFLETLQNDVCLPKATHQKEFDTLGKLESPDK 613
Db 529 EHDLEVASSEBQERESENNQPVVEERKKHNNEMEVSAHNIH-----DGATDDAEDDDDD 584
QY 614 DGLLK-----PTCGMKISLPNKALELKORETFAEDVSSVESTFS-LFGKPTT 660
Db 595 DGLIKRKSGETDHOQFPRKENKEYASGPAQOMKEVSKTEKERTSKESVNSPVFGKASL 644
QY 661 ENSQSTKVEEDFNLTKEGATKTVTGQOERDIGIIERAPODQTNKMPSELGRKEDTKST 720
Db 645 LTGGLQVDDDDSSLS-----EIDEDEGRPTKTSNEKNKV-----KNQIQSM 686
QY 721 SD-SLIIISVSDTONYEC-LPEATYQKEIKTTNGKIEESPERKSHFEPATEMQNSVPNKGL 778
Db 687 DDDVDDLTOSSETASEDCPLHSSYK-----NFMLLIQOLGM 722
QY 779 EWNKQTLRADSTTLKILDALPSCERGRELKONCEQITAMOMKMKVFCVLOKELSEA 838
Db 723 ECK-----DSVSLKKIQAALSCERLLEKONCELLTVKIKKMDKVNVLQRELSSET 775
QY 839 KEIKSOLENOKAKWEQELCSVRLPLNQEERKRNVDILKEKI-----RPEOLRKKLEVK 893
Db 776 KEIKSOLSHQKWEVERELCSLRFSLNQEERKRNADTLVEKIREQLRRKEEQYRKEVEV 835
QY 894 HOLEOTLRIQDIELKSVTSNLNOV----- 917
Db 836 QOLESLQTLMELELTVKSLNNOVQERNDQAQOLRSQNAARMQDGLITNLHLSKQKEIE 895
QY 918 -----SHTHESENDLPHENCMKKEIAMLKLEVALTKHQHQVVENKYFEDIKILQ 967
Db 896 MAQKMNSENSHSEESKDLGSHKMSLQBEIAMLRLDITIKNQONQEKKEKCFDLKIVK 955
QY 968 EKNALQWLTKLQKTVTKRASOYREQLKVLTAENTMLTSLK 1010
Db 956 EKXEDLQTKIKONEETLTQTIISQYNGRSLVLTAEANMLNSKLE 998

RESULT 5
QH1Q1 ID Q9H1Q1 PRELIMINARY; PRT; 1710 AA.
AC Q9H1Q1;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE BA145E8.1 (KIAA1074).
GN Name=BA145E8.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 4.
DR PRINTS; PR01415; ANKYRN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
```

```
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1710 AA; 196409 MW; 01CBF9BADB894872 CRC64;

Query Match 17.8%; Score 920.5; DB 2; Length 1710;
Best local similarity 28.5%; Pred. No. 5.2e-28;
Matches 320; Conservative 158; Mismatches 341; Indels 305; Gaps 35;

QY 7 SYGAVIEV-----NKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESKC 53
Db 55 SAGWAKVQVQILLRKNGLNDRDMKRTALHACANGHEVTVLLVDRKQCLNCDNEIR 114
QY 54 TALMLAICEGSSSIVGMQLQNVDFPAEDIHGTASYRYAARGVNIHQOLLEHRIKLPK 113
Db 115 TALMKAVOCQEEKCATILLEHGADPNLADVHGNTALHYA-----VYNEDISSVATKLL 167
QY 114 NPQNTPEGTSTGTPDEAALERTPDTAESLLEKTPDEAARLVEGTSAKIQCLGKATSG 173
Db 168 YDANIEAK-----NKODLTPL-----LL-----AVSG 189
QY 174 KFEQSTTEETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTEVL 233
Db 190 KQOMVEFLIKK-KANVNAVDKL-----ESSHQLISEYKEERI----- 226
QY 234 EKGTSNMIACTKETSTKASTNVDSVSEPIFSLFGTRTITENSQCCTKVEEDFNLAIIIS 293
Db 227 -----PKHSSQNSNVDESSEDSLSLGGKPGVDSDPTSDDEDLNFDTKNVP 274
QY 294 K-----SAAQNYTCLPDATYQDKITINHKI-EDQMPFSESKREDEEYSDSGSLPE 345
Db 275 KPSLAKLMTASQOSRKNLEATY-GTVRTGNRTLFED-----RSDSQDEVVVESLPT 325
QY 346 SSATQVQICIPESMVQ-----KWEINRREVEELPEKPSAFKP-AVEMOKTVPNKA 393
Db 326 TSIKQV-CFSPHYQSPDLLPKPSHSLANPLGMKSEPTKPGAKKENGIDIESAP--- 381
QY 394 FELKNEOTL-----RAAOMFPSESOKODDENSWSESPCTVSKQDVLPKATH 443
Db 382 LEQTNNDNLTVYDVHKNRDMMSALGLGQEEIDESPWSESIENFPQKYV----- 434
QY 444 QKEFDTLGKLESPVKGDLKPTCGRKVSLPNKALELKORETFAEDVSSVESTFS-LFGKPT 503
Db 435 ----DPLAGAAD-----GKEKNIGNEQAE-----DVFYIPSC- 462
QY 504 RKVSLPNKALELKORETL-----KAESPNDGLKPTCGRKVSLPNKALELKORETPKAA 558
Db 463 MSGSRNFMAKLEDTNRVGMVAHMESEPERYLHLKPTIEMKDSVPNKAGGKMDVQTSKAA 522
QY 559 Q-----MFPSESOKODDENSWDFESFLETLQNDVCLPKATHQKEFDTLGKLESPDK 613
Db 523 EHDLEVASSEBQERESENNQPVVEERKKHNNEMEVSAHNIH-----DGATDDAEDDDDD 578
QY 614 DGLLKPTCGMKI-----SLPNK-----ALELKORETFAEDVSSVESTFS-LFGKPT 659
Db 579 DGLIKRKSGETDHOQFPRKENKEYASGPAQOMKEVSKTEKERTSKESVNSPVFGKAS 638
QY 660 TENSQSTKVEEDFNLTKEGATKTVTGQOERDIGIIERAPODQTNKMPSELGRKEDTKS 719
Db 639 LLTGGLQVDDDDSSLS-----EIDEDEGRPTKTSNEKNKV-----KNQIQS 680
QY 720 TSD-SLIIISVSDTONYEC-LPEATYQKEIKTTNGKIEESPERKSHFEPATEMQNSVPNK 777
Db 681 MDDVDDLTQSSSETASEDCPLHSSYK-----NFMLLIQOLG 716
QY 778 LEWNKQTLRADSTTLKILDALPSCERGRELKONCEQITAKMOMKMKVFCVLOKELSEA 837
Db 717 MECK-----DSVSLKKIQAALSCERLLEKONCELLTVKIKKMDKVNVLQRELS 769
QY 838 AKEIKSOLENOKAKWEQELCSVRLPLNQEERKRNVDILKEKI-----RPEOLRKKLEVK 892
Db 770 TKEIKSOLSHQKWEVERELCSLRFSLNQEERKRNADTLVEKIREQLRRKEEQYRKEVEV 829
QY 893 KHOLEOTLRIQDIELKSVTSNLNOV----- 917
```

```

Db 830 KQOLESQTLMEURTVKSNLNQVQBRNDAQRLSREQNAQLDGLTNHLSKQKEI 889
Qy 918 -----SHTSESENDLFHENCMLKKEIATMLKLEAVATLKHOHQVKENKYPEDIKIL 966
Db 890 EMAQCKMNSNSHSEEEKDLSKNSMLQOEIAMLRLIEDTIKQNOKEKCCPEDLKIV 949
Qy 967 QEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKJK 1010
Db 950 KEKNEDLQTKIQNEETLTQTISQYNGRLSVLTAEANMLNSKLE 993

RESULT 6
Q6ZR14 PRELIMINARY; PRT; 1080 AA.
AC Q6ZR14;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46736.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Trachea;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128577; BAC87508.1; -.
FT NON-TER 1080
SQ SEQUENCE 1080 AA; 126741 MW; 47BBA933AE595C CRC64;

Query Match 13.6%; Score 703; DB 2; Length 1080;
Best Local Similarity 32.7%; Pred. No. 9.8e-20;
Matches 197; Conservative 98; Mismatches 163; Indels 144; Gaps 17;

Qy 485 ETFKAESPDKGLKPTCGRKVSLPNKALELKDRETLKAESPDPDGLLKPTCGRKVSLPN 544
Db 17 DAYKTRPIQLNLFQKPLYDHCSANNYKSMPELENVRSSPRGDT-----SKVSLKE 69
Qy 545 KALELKURETFK-AAQMPPESSKQKDEENSWDFESFLETLQNDVCLPKATHQKE----- 599
Db 70 ELQO--DMQRFKNRIGMLKVFEQALEKE-----KVQLQKEVEEERKKHNNMEV 117
Qy 600 ----FDTLGKLEESPDKGLKPTCGMKI---SLPNK-----ALELKURETFKA 642
Db 118 SANTHDGATDAEDDDDDGLIQRKSGETHQOPFRKENKEYASSGPALQMKVKEKTEK 177
Qy 643 EDVSSVESTFS-LFGKPTTENSQSTKVEEDFNLTKEGATKVTGQOERDIGIIERAPOD 701
Db 178 EKRTSKESVNSPVFGKASLLTGLQLQVDDSSLS-----EIDEDSGRPTKTSN 226
Qy 702 QTNKMPSTSELGRKEDTKSTSD-SHIIISVSDTQNVCE-LPEATYQKEIKTNGKIESEPK 759
Db 227 EKNKV-----KNQIQSMDDVDLDTOSSETASEDCELPHSSYK----- 263
Qy 760 PSHPFEPATEMNSVPNGKLEWKNQTLRADSTTLISKILDALPSCGERELKDKNCEOITA 819
Db 264 -----NFMLLTEQLGWECK-----DSVSLKIQDAALSCERLELKNHCELLTV 308
Qy 820 KMEQKMKFVQLKELSEBAKIKSLENQKAKWEQLCSVRLPLNQBEERKRNVDILKEK 879
Db 309 KIKKMKDVNVLQRELSETKEIKSLEHQKVEWERELCSLRFSLNQBEERKRNADTLVEK 368
Qy 880 I-----RPEQLRKLEVKHOLEOTLRIQDIELKSVTSNINQV----- 917

```

```

Db 369 IREQLRRKEBOYRKEVEVKQOLESQTLMEURTVKSNLNQVQBRNDAQRLSREQNA 428
Qy 918 -----SHTSESENDLFHENCMLKKEIATMLKLEAVATLKHOHQVKENKYPEDIKIL 948
Db 429 RMLQDGLTNHLSKQKEIEMAQCKMNSNSHSEEEKDLSKNSMLQOEIAMLRLIEDTI 488
Qy 949 KHQHVKNKYPEDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSK 1008
Db 489 KNQNOKEKCCPEDLKIVKEKNEDLQTKIQNEETLTQTISQYNGRLSVLTAEANMLNSK 548
Qy 1009 LK 1010
Db 549 LE 550

RESULT 7
Q8N7B8 PRELIMINARY; PRT; 168 AA.
AC Q8N7B8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ25822.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK098688; BAC05379.1; -.
SQ SEQUENCE 168 AA; 18305 MW; BC6BA50DB58944CF CRC64;

Query Match 10.7%; Score 553; DB 2; Length 168;
Best Local Similarity 73.9%; Pred. No. 8.6e-15;
Matches 113; Conservative 9; Mismatches 27; Indels 4; Gaps 1;

Qy 142 AESLLEKTPDEAARLVVEGTSKIQCLGKATSGKFEQSTETPRKILRPTKETSSEKFWPA 201
Db 2 AESLVERPDEAAPLVVEGTADKIQCLGKATSGKFEQSAEETPKIMRTAKETSSEKFWPA 61
Qy 202 KERSRKLTWEKETSVKTECVAGVTPNKTEVLEKGTSMNIACPTKETSTKASTNVDSV 261
Db 62 KERPRKITWEKETSVKTECVAGVTPNKTEVLEKGTSEMLTCTPTKETSTKASTNVDSV 121
Qy 262 EPIPSLF-----GTRTIENSQCTKVEEDFNLATK 290
Db 122 ESIFRVSPCHPGWNAMARSRLTATSTSQKEATK 154

RESULT 8
Q9HCD1 PRELIMINARY; PRT; 718 AA.
AC Q9HCD1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE KIAA1641 protein (Fragment).
GN Name=KIAA1641;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;

```

RT "Prediction of the coding sequences of unidentified human genes.
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which
 RT code for large proteins in vitro."
 RL DNA Res. 7:273-281(2000).
 DR EMBL; AB046861; BAB13467.1; -.
 FT NON TER 1
 SQ SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;

Query Match 9.5%; Score 489; DB 2; Length 718;
 Best Local Similarity 28.5%; Pred. No. 1.4e-11;
 Matches 160; Conservative 79; Mismatches 186; Indels 136; Gaps 12;
 Qy 457 SPVKGGLKPTCGRKVSLPNKALELKDETFKAEKSPDKGLLPTCGRKVSLPNKALELK 516
 Db 11 SSQKQPAKATSKDKDSVNPTEIKDQCGSGTSSQKQPAWKAATSVKDSVSNIAIEIK 70
 Qy 517 DRETLKAEKSPDGLLPTCGRKVSLPNKALELKDETFKAAQMPFSESQKQDEENSWD 576
 Db 71 DQIRGTVPQKSAQKVFKKKVSLLNATRI 108
 Qy 577 FESFLETLQNDVCLPKATHQKEFTLSGKLEESPDGGLKPTCGMKISLPNKALELK 636
 Db 109 -----SG--TEYPENLPTLKATIENTKNSVLNATKWK 139
 Qy 637 RETFKAEDVSVSTFSLFGKPTTENSQSTKVEDFNLTKEGATKTVTGQOERDIGIIE 696
 Db 140 -----VQTSPTQDLEWASE-----GQKR----- 159
 Qy 697 RAPQDTNMPSTSELGRKEDTKSTD-SEIISVSDTQNYECLPEATYQKEIKTTNGKIEE 755
 Db 160 --LEBYENNPQGV---KNQIHSRDDLDIIQSSQTVS----- 191
 Qy 756 SPEKSHPEPATQMSVNPNGLEWKNKQTLRADSTTSLKILDALPSCERGRELKKNCE 815
 Db 192 -----EDGSLCCNCKNVILLIDQHEMKCKDCVHLKLTNTFCLWKRLIKLDNHCE 243
 Qy 816 QITARMEQWKNKFCVLOKELSEAKIKSQLENQAKWEQELCSVRLPLNQEEKRRNVDI 875
 Db 244 QLRVKIRLKNKASVLQRISEKEIKSQLEKLELEKELCSLRFALQEKKKRRNVEE 303
 Qy 876 LKEKIR-----PEQRLKLEVKHOLEQTLRIQDIELKSVTSNLSNVSHSESENDLPE 930
 Db 304 LHQKREKLRIITEQYRIEADVTKEIPKALKAESAELTKTGGNSNQVSETDEKE--DLLHE 362
 Qy 931 NCMLKKEATMLKLEVALTKHGVENKVFEDIKILOEKNAELOMTLKLKQKTVTKRASQ 990
 Db 363 NRMQDEIARLLEKDTIKNQ--LEKKYLKDFEIVKVKHEDLOKALRNGETLAKTIAC 420
 Qy 991 YREQLKVLTAENTMLTSKIKE 1011
 Db 421 YSGQLAALTDTENTLRSKLEK 441

RESULT 9

Q992S2 ID Q692S2 PRELIMINARY; PRT; 1043 AA.
 AC Q692S2
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DE MKIAA1074 protein (Fragment).
 GN Name=MKIAA1074;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
 RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
 RA Nagase T., Ohara O., Koga H.;
 RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
 IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous

RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
 RT Randomly Sampled from Size-Fractionated Libraries."
 RL DNA Res. 11:205-218(2004).
 DR EMBL; AK173096; BAD32374.1; -.
 DR InterPro; IPR010989; c-snare.
 DR InterPro; IPR009054; Topismrse_insert.
 FT NON TER 1
 SQ SEQUENCE 1043 AA; 121354 MW; 5F37B1F9AA37B212 CRC64;

Query Match 9.0%; Score 465; DB 2; Length 1043;
 Best Local Similarity 33.9%; Pred. No. 1.9e-10;
 Matches 122; Conservative 56; Mismatches 84; Indels 98; Gaps 9;
 Qy 701 DQNKMPSTSELGRKEDTKSTDSEIISVSDTQNYECLPEATYQKEIKTTNGKIEESPEKP 760
 Db 91 DQSEAPT-----KTSKKNKDS-----GQTAADVNLDDTFESETASEDHDL-QGPDSE 140
 Qy 761 SHFEPATEMNSVNPNGLEWKNKQTLRADSTTSLKILDALPSCERGRELKKNCEQITAK 820
 Db 141 S-----ILCAIEHLRLCK-----DTASLLKIRDVAVSYKRLIELKRSHCELLTGK 186
 Qy 821 MEQWKNKFCVLOKELSEAKIKSQLENQAKWEQELCSVRLPLNQEEKRRNVDIKKEKI 880
 Db 187 LKRMENKYKGLQKEMSGTEEVKSRLEBHEKVGWQELCLRLFALKQEEKKRSADQLSEKT 246
 Qy 881 -----RPEQRLKLEVKHOLEQTLRIQDIELKSVTSNLSNV----- 917
 Db 247 MEQLRRKGEOCQSEVEARQQLSEASRLTLEMLATVKVSHLNQVLEERNETQRLSREQNAR 306
 Qy 918 -----SH-----THESENDLFHENCMLKKEIAMLKLEVATLK 949
 Db 307 MLQDGLLASHLCKQKEIEMTQKMTSEVSVSHEKEKDLKHQRLQDEVAVLREMDTIK 366
 Qy 950 HQHQVKENKVFEDIKILOEKNAELOMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKL 1009
 Db 367 SHNQEKERYLEDIKIANEKNDLNQVLU-----NMLSSKL 403

RESULT 10

Q9H0H6 ID Q9H0H6 PRELIMINARY; PRT; 823 AA.
 AC Q9H0H6
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE Hypothetical protein DKFZp434A171.
 GN Name=DKFZp434A171;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RG The German cDNA Consortium;
 RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
 RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL136793; CAB66727.1; -.
 DR HSSP; Q60778; 10Y3.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 6.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 6.
 DR PROSITE; PS0088; ANK_REPEAT; 4.
 DR PROSITE; PS0297; ANK_REPEAT; 1.
 KW ANK repeat; Hypothetical protein.
 SQ SEQUENCE 823 AA; 93968 MW; 9D2DD5F52C379B2 CRC64;

Query Match 7.7%; Score 396; DB 2; Length 823;
 Best Local Similarity 17.6%; Pred. No. 7.2e-08;
 Matches 178; Conservative 100; Mismatches 155; Indels 580; Gaps 18;

Qy	5	LLSYGAVIEVQNKASLTPLLLAIQKRSQOTVEFLTTKUANANAFNESCCTALMLAI	CEGS	64
Db	152	LLSHGAHTEALDKONNTPLFAIICKKQWVEFLKKKASSHAVDRLRSSALMLAVYDS	211	
Qy	65	SEIVGMLLOQNVDPFAEDIHGITAERYAARGVYIHQOLLEHPIKLPQNPQNTWPGTS	124	
Db	212	PGIWNILUKQINDVFAQCMGRDAEYAISSHHLTKIOQOILEHKKILKK-----	EKSD	265
Qy	125	TGTPDEAAPLAERTPDTAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFPQSTEPR	184	
Db	266	VGSSDESA-----	-----	273
Qy	185	KILRPTKETSEKPSWPAKERSRKITWBEKETSVKTECVAGVTPNKTVEULEKTSNNIACP	244	
Db	274	-----	-----	273
Qy	245	TKETSTKASTNVDSVSSVEPIPSLFGTRTIENSQCTKVBEEDFNLATKIISKSAANYTCLP	304	
Db	274	-----VSIFHELVDSLPSAD-DKOLNVATK-----	QCVP	302
Qy	305	DATYQKDIKTINHKIEDQMPFSESKEEDBYSDSGSLFESSAKTQVCIPESMYQKVM	364	
Db	303	E-----KVSPELPSSHEKGNR	319	
Qy	365	I-NREVEELPEKPSAFKPAVEMQKTPVPNKAFELKNEQTLRAAQMPFSSKQXDDENSWD	423	
Db	320	IVNGQEGEPKAKHPSLKPSTEVEDPAVKGAQVRKNVQTLRAEQALPVASEEBOER----	374	
Qy	424	SESPCETVSQKDVLPKATHQKEFPTLSGLKEESPVKDGLLKPTCGRKVSLPNKALELKD	483	
Db	375	-----HER-----	-----	377
Qy	484	RETFKAESPDKGLLKPTCGRKVSLPNKALELKDRETLKAESPNDNDGLLKPTCGRKVSLP	543	
Db	378	-----	-----	377
Qy	544	NKALELKDRETFKAAQMPFSESQKQDDENSWDFSEFLETLQNDVCLPKATHQKEFOTL	603	
Db	378	-----SEKKQPVQKEGN-----	-----	389
Qy	604	SGKLEESPDKGLLKPTCGMKLSLPNKALELKDRETFKAEDVSVSESTFSLFGKPTTENS	663	
Db	390	-----NTNKS	394	
Qy	664	QSTKVEEDFNLTKEGATKVTGQOERDIGIIERAPQDOTNKMPTSELGRKEDTKSTDSDS	723	
Db	395	EKIQLSENCIDSTSSAAAGRLT--QORKIG-----	-----	422
Qy	724	EIISVDSQNYECLPEATYQKEIKTTNGKIEBSEPEKPSHFEPATEMQNSVPNKGLEWKNK	783	
Db	423	-----KTY-----PQFP-----	-----	430
Qy	784	QTLRADSTTSLKILDALPSCBGRGLKQDNCCEQITAKMEQMKPCVQLQKELSEAKEIKS	843	
Db	431	-----	-----	430
Qy	844	QLENQAKWEQELCSVRLPLNQEBEKKRNVDL-----KEKIRPEQLRKKLKVHQLBQ	898	
Db	431	-----KKLKEHRDCT-----LKQNESEKTNVNMLYKKNREELERKEKQYKVEAK-QLEP	481	
Qy	899	TLRIQDLEKSVT--SNLNOVSHTESENDLPHENCMKKETIAMLKLEVALTLKHQHVKE	956	
Db	482	T--VQSLEMKSKTARNTNRDFHNHEEMKGLMDENCILKADIAIRQELICTWMDNLEKE	539	
Qy	957	NKYFEDIKILQFKNAELQWTLKQKTTVTKRASQYREQKVLTAENTMLTSKL	1009	
Db	540	NKYLKDIKIVKETNAALEKYIKLNEEMITETAFRYQQLNDLKAENLRNAEL	592	
RESULT 11				
Q7K5Q6				
ID	Q7K5Q6	PRELIMINARY:	PRT: 2006 AA.	

AC	Q7K5QG;
DT	05-JULI-2004 (TrEMBLrel. 27, Created)
DI	05-JULI-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JULI-2004 (TrEMBLrel. 27, Last annotation update)
DE	Erythrocyte binding protein 3.
CN	Name=aebl;
OS	Plasmodium falciparum.
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=5833;
RN	[1]
RP	SEQUENCE FROM N.A.
EX	MEDLINE=22072559; PubMed=12076768; DOI=10.1016/S0166-6851(02)00067-1;
RA	Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
RT	"Plasmodium falciparum MAEBL is a unique member of the ebl family.";
RL	Mol. Biochem. Parasitol. 122:35-44(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
EX	MEDLINE=22077637; PubMed=12082132;
RA	Nichon P., Stevens J.R., Kaneko O., Adams J.H.;
RT	"Evolutionary relationships of conserved cysteine-rich motifs in
RT	adhesive molecules of malaria parasites.";
RL	Mol. Biol. Evol. 19:1128-1142(2002).
RN	[3]
RP	SEQUENCE FROM N.A.
RA	Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY042084; AAQ73469.1.;
DR	InterPro: IPR008602; Duffy binding.
DR	Pfam: PF05424; Duffy-binding. 1.
SQ	SEQUENCE 206 AA; 237620 MW; EE75EP42E2BF76D CRC64;

Query Match 6.9%; Score 359; DB 2; Length 2006;
 Best Local Similarity 22.4%; Pred. No. 5.6e-06;
 Matches 233; Conservative 159; Mismatches 398; Indels 248; Gaps 42;

Qy	109	RKLPRNQNTNPETGTTGTDEAAPLAERTPDTAESLLEKTPDPAARLVGSTSAKIQCIG	168
Dd	828	RNLSPAYNHLTNECVILGTHEQ-----ERTNSCRTKEEK-----KKPNCQLR	871
Qy	169	KATSGK---PQSSTPTPKLRITKTSEKFSPAKERSRKI--TWEEK-----	213
Dd	872	KTTDSKDWTYVSS-----FIRDPYTKCPRPPLKS---KVGTFDQTKCKCSLMXK	921
Qy	214	-----ETSVKTECVAGVPINPTEVLKGTSGNNIACPRTKETSKASTNVDVSS-----	260
Dd	922	AYEVGINFKSVCLVFLVSPK---DLYNSGRNNYWGIIWAADHSYN-ENNIEANGKCVMH	978
Qy	261	-VEPIPSLFGTRTTENSQCTKVEDDFNLATKISKSAQNVTCLPDATYQDKYT----	314
Dd	979	VVKPTCIDKENHFSTALTANTVDFFQSNVIRIKIBELTXYGNDDVLKKEKINPEIDN	1038
Qy	315	-----INHK-----IEDQMFPESKREDEBSWDGSLSPESSAKTV--CIP	355
Dd	1039	VNETKINKSHVNMERNKPSYKENYDQMEKNVEDETYSSEFG-LFEARKTETGRLEE	1097
Qy	356	ESMTQKYW-----EINREVE-----E 371	
Dd	1098	ESKKKAMKRAEDARRIEEARAEADARRIEARRAEDARRVEIARRVEDARRIEISRAE	1157
Qy	372	LPEKPSAFKPAVENMQKTVPNKAFLKNEQTLLRAAQMPFSKQKDDSENSWSSPCETV	431
Dd	1158	DAKRIEAARAIAEVRRAELRKA---EDARRIEAARY--ENERIEEARRYEDEKRIEAV	1212
Qy	432	S-----QKDVYLPLKATHQKPFDTLSGLKESPV-----KOGLLKPTCGRVSLPNKALEL	481
Dd	1213	KRAEVRKDBEEAKRAEKERNNEIRKFEPEARMHAFAHQRAAIKAEKRADELKKAEEK	1272
Qy	482	KDRETFFKAESPDK--DGILLKPTCGRVSY--LPNKALELKORETLKABSPNDGLLKPTCG	537
Dd	1273	KAADELKKESEKKAADELKKAEEKKAADELKKAEEKKAADELKKAEEKK----KADEV	1329
Qy	538	RVVSLPNKALEL-KDRETFKAAQMPFSKQKDDSENSWDFSETLTLQNDVCLPKATH	596

RESULT 11
Q7K5Q6
ID Q7K5

Db 1330 KKAEEKKADELKKSEKKKKADELKSKSEKKKADE-----LKKAAEKKKADELKKAAE 1384
 Qy 597 QKEPDTLSGKLESPDKGLLKPTCGMKISLPNKALELKDR--ETFFKAEDVSSVESTFSL 654
 Db 1385 KKADELKKKAAEKKKADELKK-----KAAEKKKADELKKKAAEKKKADELKKKAAE---1436
 Qy 655 FGKPTTENSQSQTVEEDFNLTATTKGATKTVTGQOQERDIGIIERAPQDOTNKMPTSELGRK 714
 Db 1437 --KKAENLK--KAE-----KKKADELKKAAEKKKADELKKAAE-----KKKADELKKK 1484
 Qy 715 EDTKSTSDSEIISVSDTQNYECLPRTATYQKEIKTTNGKIEBSPKPSHFHPATQMNSVP 774
 Db 1485 AEEKKAD-ELKKAEEKKKADELKKAAEKK--KADELKKAAEKKKADELKKAAELKKAEE 1541
 Qy 775 NKGLFWNKQTLRADSTTLTKILDALPSCERGR-----ELKKNCEQI 817
 Db 1542 KKKVEQKKREBERNNAL--RRAELKQIEKKRIEVMKLYEEKKKMAEQKKKEBEKI 1599
 Qy 818 TA-----KMEQMKNFVLOKSELSEAKIKSQLENQKAKWQELCSVRLPLNQE 866
 Db 1600 KAEQKKKEBEKKVEQKKK--ESEEKKKAEQKKKEEENKIKAEQ-----LKKKEE 1650
 Qy 867 EKKRNVDILK---EKIRPEEQRLKLELVHQHLEQTLRIQDIELKSVTSLNQVSHTHES 923
 Db 1651 EKKKAELKKEEKKKAEQKKKEBEKKKVEQKKKVEQKKKKEEKKKAE-----1698
 Qy 924 ENDLPHENCMKKIEIAMLKLEVATLKHQHVKENKYPEDIKLOEKNAELQMTLKLKQKT 983
 Db 1699 -----QKKKEEENKIKVEQKKKEE--EKKKAELKKEEKKKVEQKKKKEE--1746
 Qy 984 VTKRASQYREQLKVLTAE 1001
 Db 1747 --KKAEEIRKEAEVIEE 1762

RESULT 12
 ID Q7K5Q5 PRELIMINARY; PRT; 2019 AA.
 AC Q7K5Q5;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE Erythrocyte binding protein 2.
 GN Name=maebi;
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_taxid=5833;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22072559; PubMed=12067678; DOI=10.1016/S0166-6851(02)00067-1;
 RA Blair P.L., Kappe S.H., Maciel J.E., Balu D.J., Adams B.;
 RA Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
 RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
 RL Mol. Biochem. Parasitol. 122:35-44(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22077637; PubMed=12082132;
 RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
 RT "Evolutionary relationships of conserved cysteine-rich motifs in
 RT adhesive molecules of malaria parasites.";
 RL Mol. Biol. Evol. 19:1128-1142(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY042084; AAQ73469.1; -;
 DR InterPro; IPR008602; Duffy binding.
 DR Pfam; PF05424; Duffy binding; 1.
 SQ SEQUENCE 2019 AA; A626F2C684C08785 CRC64;

Query Match 6.9%; Score 359; DB 2; Length 2019;
 Best Local Similarity 22.4%; Pred. No. 5.6e-06;
 Matches 233; Conservative 159; Mismatches 398; Indels 248; Gaps 42;

Qy 109 RKLPNQNTNPECTSTGTDEAPLAERTPDTAESLLEKTPDEARLVEGTSAKIOCLG 168
 Db 828 RNLPAYNHLTNECVILGTHEQ-----ERTNSCRRTKEE-----KKNCQILR 871
 Qy 169 KATSGK---EQSTTEETPRKILRPTKETSEKFSWPAKERSRKI---TWEEK- 213
 Db 872 KTTDSKDMTWYSS-----FIRPDYETKCPPRYPLKS---KVFGTDFDKTGCKSLMDK 921
 Qy 214 -----ETSVKTECVAGVTPNKTEVLEKGTSMIACTKETSTKASTNVDDVS-----260
 Db 922 AYEVGINKFSVCLYLEFLVSPK--DLYNSGRNNWGIWAADHSVN--ENNIANGKCVHL 978
 Qy 261 -VEPIFLSGFTRITENSQCTKVEEDFNLTATKIISKSAQNYTCLDPATYQKDIKT-----314
 Db 979 VVKPTCVIDKENHFSFTALTANTVDFNQSVNIRKIEELTEYGNNDVVLKEKEINNEPIDN 1038
 Qy 315 -----INH-----IEDQMPSESKEDEEYSDWSGSLPESAKTOV--CIP 355
 Db 1039 VNETHKINRSHVNSMERKNKPSYKENYEDOMEKNVEDETYSEFG-LFEEARKTGTGRIE 1097
 Qy 356 ESMYQKVM-----EINREVE-----E 371
 Db 1098 ESKKEAMKRAEDARRIEEARRAEDARRIEEARRAEDARRIEEARRAEDARRIEEARRA 1157
 Qy 372 LPEKPSAFKPAVENQKTPVNNKAFELKNEQTLRAAQMPFSESQKDDSENSWDSSEPCETV 431
 Db 1158 DAKRIEARRAIEVRRALRKA---EDARRIEAARY--ENERRIEABARRYEDEKRIEAV 1212
 Qy 432 S-----QKDVYLPATHQKEPDTLSGKLESPV-----KQGLLKPTGCRKVSLEPNKALEL 481
 Db 1213 KRAEEVRKDBEEAKRAEKKERNNEIRKFEEMAHFARRQAAIKAEKRADELLKKAEEK 1272
 Qy 482 KDRETFKAESPDK--DGLLKPTCGRKVS--LPNKALELKDRTELKAEKSPDNDGLKPTCG 537
 Db 1273 KKADELKSEKKKADELKKAAEKKKADELKKAAEKKKADELKKAAEKK--KADEV 1329
 Qy 538 RVSLEPNKALEL-KDRETFKAAQMPFSESQKDDSENSWDFSEFLETLQNDVCLPKATH 596
 Db 1330 KKAEEKKKADELKKSEKKKADELKKSEKKKADE-----LKKAAEKKKADELKKAAE 1384
 Qy 597 QKEPDTLSGKLESPDKGLLKPTCGMKISLPNKALELKDR--ETFFKAEDVSSVESTFSL 654
 Db 1385 KKADELKKKAAEKKKADELKK-----KABEKKKADELKKKAAEKKKADELKKAAE---1436
 Qy 655 FGKPTTENSQSQTVEEDFNLTGATKTVTGQOQERDIGIIERAPQDOTNKMPTSELGRK 714
 Db 1437 --KKAENLK--KAE-----KKADELKKAAEKKKADELKKAAE-----KKADELKKK 1484
 Qy 715 EDTKSTSDSEIISVSDTQNYECLPRTATYQKEIKTTNGKIEBSPKPSHFHPATQMNSVP 774
 Db 1485 AEEKKAD-ELKKAEEKKKADELKKAAEKK--KADELKKAAEKKKADELKKAAELKKAEE 1541
 Qy 775 NKGLFWNKQTLRADSTTLTKILDALPSCERGR-----ELKKNCEQI 817
 Db 1542 KKKVEQKKREBERNNAL--RRAELKQIEKKRIEVMKLYEEKKKMAEQKKKEBEKI 1599
 Qy 818 TA-----KMEQMKNFVLOKSELSEAKIKSQLENQKAKWQELCSVRLPLNQE 866
 Db 1600 KAEQKKKEBEKKVEQKKK--ESEEKKKAEQKKKEEENKIKAEQ-----LKKKEE 1650
 Qy 867 EKKRNVDILK---EKIRPEEQRLKLELVHQHLEQTLRIQDIELKSVTSLNQVSHTHES 923
 Db 1651 EKKKAELKKEEKKKAEQKKKEBEKKKVEQKKKVEQKKKKEEKKKAE-----1698
 Qy 924 ENDLPHENCMKKIEIAMLKLEVATLKHQHVKENKYPEDIKLOEKNAELQMTLKLKQKT 983
 Db 1699 -----QKKKEEENKIKVEQKKKEE--EKKKAELKKEEKKKVEQKKKVEQKKKKEE--1746
 Qy 984 VTKRASQYREQLKVLTAE 1001
 Db 1747 --KKAEEIRKEAEVIEE 1762

RA Ghai M., Dutta S., Hall T., Freilich D., Ockenhouse C.;
RT "Identification, expression, and functional characterization of MAEBL,
RT a sporozoite and asexual blood stage chimeric erythrocyte-binding
RL protein of Plasmodium falciparum";
DR Mol. Biochem. Parasitol. 123:35-45 (2002).
DR EMBL; AF400002; AAM90625.1; -;
DR HSP; P04368; IIC2.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 2055 AA; 243211 MW; 783F48BB7286B1F3 CRC64;

Query Match 6.9%; Score 359; DB 2; Length 2055;
Best Local Similarity 22.4%; Pred. No. 5.7e-06;
Matches 233; Conservative 159; Mismatches 398; Indels 248; Gaps 42;

QY 109 RKLKPNPONTNPEGSTGTGTPDEAALPABRTPTDAESLLEKTPDEAARLVEGTSAKIOCLG 168
DB 828 RNLPAYNHLNNECVILTHQ-----ERTNSCRTKEE-----KKPNCQLLR 871
QY 169 KATSGK---FQSTETPRKILRPTKETSEKFSWPAKERSKI--TWEEK----- 213
DB 872 KTTDSKOWTVYSS-----FIRPDYETKPPRYPLKS--KVFGTDPQTKGCKSLMDK 921
QY 214 -----ETSVKTECVAGVTNKTETVELEKGTNSMIACPTKSTYKASTNVDS-- 260
DB 922 AYEVGINKFSFLEVLFLVSPK--DLYNSGRNYYGWIWAADHSVN--ENNBIANGKCYHL 978
QY 261 -VEPIFSFGRTIENSQCTVEEDFNATKIISKSAQNTCLPDTAYQDKT----- 314
DB 979 VVKTCVIDKENHFSFTALTNTVFNOSVIRKIEELTEGNDDVLKEIINNEPIDN 1038
QY 315 -----INHK-----IEDQMFPSKREDEEYSDSGSLFPSSAKTQV--CIP 355
DB 1039 VNETKINKSHVSMERNKPSYKENEYDQMEKNVEDYSEFG-LFEEARKTETGRIEE 1097
QY 356 ESMYQKVM-----EINREVE-----E 371
DB 1098 ESKKEMKRAEDARRIEBARRAEDARRIEARRAEDARRIEARRAEDARRIEARRA 1157
QY 372 LPEKPSAFKPAVEMOKTPNKAFLKNEQTIRAAQMPFSESKOKDEENSDSPCSTV 431
DB 1158 DAKRIEARRAIEVRAELRKA--EDARRIEARRY--ENERRIEARRYEDEKRIEAV 1212
QY 432 S-----QKDVLPKATHQKEFDTLGKLEESPV-----KGLLKPTCGRKVSLPNKALEL 481
DB 1213 KRAEVRKDEEKAERKERNNEEIRKFEARMMAHFARRQAAIKAEKKKADLKKAEK 1272
QY 482 KDRETFKAESPDK--DGLLKPTCGKRVK--LPNKALELKDRETLKAESPDNDGLLKPTCG 537
DB 1273 KKADELKSEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKKADLKKAEKKK--KADEV 1329
QY 538 RKVSLPNKALEL-KDRETFKAQMPFSESKOKDEENSDPFLETLTLLQNDVCLPKATH 596
DB 1330 KKAEEKKKADLKKAEKKKADLKKSEKKKADLKKAEKKKADLKKAEKKKADLKKAE 1384
QY 597 QKEFTTLGKLEESPDGLLKPTCGMKISLPNKALELKD--ETTFKAEDYSSVESTFSL 654
DB 1385 KKADELKKAEKKKADLKK-----KAEKKKADLKKAEKKKADLKKAEKKKADLKKAE 1436
QY 655 FGKPTTENSQSTKVEDENLTKGATVTVGQQRDGIIERAPQDQTNKMPSELORK 714
DB 1437 --KKAENLK--KAE-----KKKADLKKAEKKKADLKKAE-----KKKADLKK 1484
QY 715 EDTKSTSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKSPHFEPATMQNSVP 774
DB 1485 ABEKKKAD-ELKAEKKKADLKKAEKK--KADLKKAEKKKADLKKAEKKKADLKKAE 1541
QY 775 NKGLEWKNQTLRADSTTLKSLDALPSCGR-----ELKKDNCQPI 817

DB 1542 KKKVEQKKREBERNMAL--RRABILQIEIKKRIEYVMKLYEEKKMAEQLKKEBEKI 1599
QY 818 TA-----KWEOMKMKFCVLQKELSEAKIKSQLENQKAKWQELCSVRLPLNOE 866
DB 1600 KAEOLKKEBEKKVQOLKKK--ESEEKKAEOLKKEBEENKIKASQ-----LKKKEE 1650
QY 867 EEKERNVDILK--EKIRPEEQRLKLEVKHQLEQTILRIQDIELKSVTSLNQVSHTHES 923
DB 1651 EEEKKAEELKKEBEKKAEOLKKEBEKKVQOLKKEBEKKAE----- 1698
QY 924 ENDLFHNCMLKGEIAMLKLEVATLKHQHKVENKYFEDIKLOEKNAELQMTLKLQKT 983
DB 1699 -----OLKKEBEENKIKVEOLKKEE--EEKKAEELKKEBEKKVQOLKKEE-- 1746
QY 984 VTKEASQYREOLKVLTA 1001
DB 1747 --KKAEEIRKEKEAVIEE 1762

RESULT 15
Q8ISFS PRELIMINARY; PRT: 10578 AA.

ID Q8ISFS
AC Q8ISFS;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE lmda_1 protein.
GN Name:isof;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22269627; PubMed=12381307; DOI=10.1016/S0022-2836(02)00970-1;
RA Flaherty D., Gernert K., Shmeleva N., Tang X., Mercer K.,
RA Borodovsky M., Benlan G.;
RT "Titins in Caenorhabditis elegans with Unusual Features: Coiled-coil
RT Domains, Novel Regulation of Kinase Activity and Two New Possible
RT Elastic Regions";
RL J. Mol. Biol. 323:533-549 (2002).
DR EMBL; AY130758; AAM61519.1; -;
DR PIR; E89066; E89066.
DR PIR; T33247; T33247.
DR HSSP; Q10466; 1BPV.
DR InterPro; IPR008938; ARM.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR008941; TPR-like.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGE2; 3.
DR PROSITE; PS0853; FN3; 2.
DR PROSITE; PS0835; IG_LIKE; 11.
SQ SEQUENCE 10578 AA; 1174295 MW; 71BF50C502FB4F16 CRC64;

Query Match 6.9%; Score 357.5; DB 2; Length 10578;
Best Local Similarity 21.7%; Pred. No. 4.2e-05;
Matches 269; Conservative 157; Mismatches 454; Indels 357; Gaps 47;

QY 36 EFLTKNANAFNESKCTALMLAIC-----EGSSEIVG--MLQQNVDPFAEDIHGITA 88
DB 8993 EFYAFRIVAVNGFEGAPSEIIIEVNTLDYDQESFDFAGEBELKDDVQVNVVTEITI 9052
QY 89 ERYAARGVNYHQLEHRLKLPKNQNTNPEGTSTGTDPDEAAPLAERTPTDAESLLEK 148
DB 9053 EE-----SEVTIEHRKLKKKKSK--KTTDEPELDSEIALEVSSDITSSLEI 9099
QY 149 T-----PDEARLVEGTSAKTQCLGKATSGKFEQSTETPRKILRPTKETSEKFSWPAKE 203

Db 9100 TTESTIPDTPESQETLUNVEIAV-----TETTVQKINPDESASAKDVNEDTA 9147
Qy 204 RSRKITWEEKETSVKTECVAGVTPNK-----TEVLEKGT--SNMIACPTKETSTKAST 254
Db 9148 VSSIVKDDRVNKKSLPESGLTTKKEIQKPEKKIMKKTEKADSSISSETSLTKDLT 9207
Qy 255 NVDVSSVEPIPSLFGTTIENSQCTKVEEDFNATKIISK----- 294
Db 9208 QTKQSEPEP-----AKRTTE-----TSVQDEVKRTTETTSKSKQTTEBHPQGGKSDSSIS 9258
Qy 295 -----SAAQNYTCLP-----DATYQDKIK-TINHKIEDOMF- 324
Db 9259 STSDASEVKVQVQSESESAQKTEKPEAKLESKSKMTEDTTKESDNKNETVDEKPKKKVLK 9318
Qy 325 -----PSES-----KREEDERY-----SW 338
Db 9319 KTEKSPDSTISETSETSAVESAGPSESETQNVAAVDKEKKQKQKTEDEKQKLEAEIAGKKST 9378
Qy 339 DSGSLFESSAKTQVCIPESMYQKWEINREVEELPEKPSAPKPAVEMQKTVPNKAFEL-- 396
Db 9379 EQSKLEAEAKLKRAEBEDAQKQK-----KTEAASKKAABAEKLEKQAOI-NKAAEADA 9434
Qy 397 ---KNE-----QTLRAAQMPSPESKQKDDDENSWDSSEPCETVSQKDVYL 438
Db 9435 VKQNELDEQNKLEATKKLAELKLEEQSAKSKQAABEQAKLDAQTKAK-AAEKQTGL 9493
Qy 439 PKATHQKEFDTLGKLEESPVKDGLLKPTCGRKVSLPNKALELKDRETFRAESPKDKGLL 498
Db 9494 EKDEKSNKSGSNETVEEKKPKKVLKKKTEKSDSISQKS-----DTSKTVAES----- 9542
Qy 499 KPTCGRKVSLPNKALEL--KDRETLKAESPNDGLLKPTCGRKVSLPNKALELKDRETFK 556
Db 9543 ---AGSESETKVADATSKQKTDKQKLEAEITAKKSADEKSKLETES-----KLK 9593
Qy 557 AAQMFPSKSKQDDDENSWDFESFETLLQNDVCLPKATHOK-----EPDTL 603
Db 9594 AAEDAARKQKQEKED-----KUKLEADVASKKAAAEKLEKQAOIKKAAEADAV 9642
Qy 604 SGKLEESPDGDLKPTCGMKISLPNKALELKDRETFKAEDVSVSESTFSLFGKPTTE-N 662
Db 9643 K-KQKLEAEKQKLESEAAATKAAAEKLE-EOAQINKAAEADAVKQKLEDEKKNLEAN 9700
Qy 663 SQST-----KVEEDFNLTKEGATKVTGQQRDGIIEERAPQDQTNKMPSTSELGRKEDTK 718
Db 9701 KKSAAEKLEKLEESAASKQ---TVBEOAKLDAQTKETAETKOT-----GLEKDDK 9748
Qy 719 STDSBIIISVSDTONYECLPEATYQKSIKTTNGKIBES-----PEKPSHFEPAT 767
Db 9749 STKDS-----SKETVDEKPKKVLKKKTEKSDSISQKSVTSKTVVSGGPPSESET 9800
Qy 768 EMQNSVPNKGLEWKNKOTLRADSTTLTKILDALPSCERRELK-----KDNCEQ 816
Db 9801 QKVADAARKQKETEOKKLEAE-ITAKKSADEKLEAEKLEKAAAEVAAKQKQKQKDEQ 9859
Qy 817 ITAKMEQMKNFVLOKELSEAKEIKSQLENQKAKWEQELC-----SVRLPL 863
Db 9860 LKLDTEAASKKAAAEKLEKQAOIKKAAEADAVKKEKLEAEKQKLESEAAATKKAABAEKL 9919
Qy 864 NOBEKERNVDILK-EKIRPEQL---RKLEV-----KHOLEQTLRTQDIE----- 906
Db 9920 KLEQKQKDAFTASIEKQKQKLEQKLEQKLEQKLEQKLEQKLEQKLEQKLEQKLEQK 9979
Qy 907 -----LKSVT-----SNLNOVSHTHE-----SENDLPH-----EN 931
Db 9980 VDEKPKKVLKKKTEKSDSISQKSDTAKTVAESAGSDSETQKVSADKAHKKQKESDEK 10039
Qy 932 CMLKKEIAMLKLEVALTKHQHVKNKYFEDIKIQEK-----NAELOMTLKLKQ 981
Db 10040 QKLESEIAAKKSAEQSKLETAETKVKVIEDSAKQKQKQKQKQKQKQKQKQKQKQKQ 10099
Qy 982 K-----TVTKRASQVREQLKVLTAENTMLTSK 1008

Db 10100 KLESEATSKKPTSEKQKDEKTPQEKAKSENETVMTTE 10136

Search completed: February 19, 2005, 02:10:53
Job time : 131.257 secs

1115 TUGT BLINK (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:56:15 ; Search time 28.5932 Seconds
(without alignments)
3402.035 Million cell updates/sec

Title: US-09-602-362E-27

Perfect score: 5173

Sequence: 1 MVATLLSYGAVIEVQNKASL.....REQLKVLTAENTMLTSLKE 1011

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	341.5	6.6	1427	2 S22695	restin - human
2	340.5	6.6	1790	2 S67593	transport protein
3	337	6.5	1392	2 A43336	microtubule-vesicl
4	321.5	6.2	1938	2 JC5421	smooth muscle myos
5	321.5	6.2	1972	2 JC5420	smooth muscle myos
6	321	6.2	2663	1 S28261	centromere protein
7	318.5	6.2	2116	2 A26655	myosin heavy chain
8	315	6.1	1972	1 A41604	myosin heavy chain
9	315	6.1	3488	2 T34418	hypothetical prote
10	312.5	6.0	5327	2 T13564	microtubule-associ
11	311	6.0	3259	1 A56539	giantin - human
12	310	6.0	1961	1 A61231	myosin heavy chain
13	310	6.0	3225	2 I52300	giantin - human
14	307.5	5.9	2954	2 T14156	kinesin-related pr
15	302.5	5.8	3187	2 JC5837	364K Golgi complex
16	300.5	5.8	1313	2 F96673	hypothetical prote
17	300.5	5.8	2253	2 T30336	nuclear/mitotic ap
18	300	5.8	1979	1 S03166	myosin heavy chain
19	295.5	5.7	2057	2 S61477	myosin II heavy ch
20	294.5	5.7	2017	1 A36014	myosin heavy chain
21	293.5	5.7	1957	2 T38077	hypothetical colle
22	293	5.7	1939	2 T18372	repeat organellar
23	292.5	5.7	1410	1 A57013	early endosome ant
24	291.5	5.6	2677	2 A38194	desmoplakin I - hu
25	290	5.6	2139	2 T18296	myosin heavy chain
26	287.5	5.6	2331	2 T25410	hypothetical prote
27	286	5.5	993	2 S49461	synaptonemal compl
28	285	5.5	1999	1 S21801	myosin heavy chain
29	284.5	5.5	1538	2 T29095	cardiac muscle fac

30	284	5.5	1642	2 T08880	NMDA receptor-bind
31	282	5.5	1558	2 B71603	RESA-H3 antigen pf
32	281	5.4	1390	2 S51364	sperm tail-specifi
33	279.5	5.4	1875	2 S38173	myosin-like protei
34	279.5	5.4	6642	2 T29757	protein UNC-89 - C
35	279	5.4	1295	2 T24587	hypothetical prote
36	279	5.4	1738	2 T14867	interaptin - slime
37	278.5	5.4	1959	1 A33977	myosin heavy chain
38	276	5.3	1837	2 T41023	probable nuclear p
39	275.5	5.3	1992	2 A47297	myosin heavy chain
40	274	5.3	1690	2 T13030	microtubule bindin
41	273	5.3	1620	2 S61535	nucleotide-binding
42	273	5.3	1804	2 T34518	nestin - golden ha
43	273	5.3	2020	2 T21174	hypothetical prote
44	271.5	5.2	1269	2 F84730	probable myosin he
45	271.5	5.2	1727	2 T50073	myosin-like coiled

ALIGNMENTS

RESULT 1

S22695

restin - human

C:Species: Homo sapiens (man)

C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004

C:Accession: S22695; S19853

R:Bilbe, G.; Delabie, J.; Brueggen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;

EMBO J. 11, 2103-2113, 1992

A>Title: Restin: a novel intermediate filament-associated protein highly expressed in th

A:Reference number: S22695; MUID:92289675; PMID:1600942

A:Accession: S22695

A:Molecule type: mRNA

A:Residues: 1-1427 <BIL>

A:Cross-references: UNIPROT:P30622; EMBL:X64838; NID:g35998; PIDN:CAA46050.1; PID:g35999

C:Keywords: cytoskeleton

Query Match 6.6%; Score 341.5; DB 2; Length 1427;
Best Local Similarity 22.4%; Pred. No. 2.5e-07;
Matches 251; Conservative 185; Mismatches 414; Indels 271; Gaps 54;

Qy	1	MVATLLSYGAVIEVQNKASLTPLLLAIOKRSKQTVFLLTKNANANAFNSKCTALMLAI	60
Db	300	MATTSASLKRSPSASSLSMSSVASSVSSRPSRT--GLLTSTSRYARKISGTTALQAL	357
Qy	61	CEGSSEIVGMLQONVDVFAEDIHGITAERYAARGVNYIHQOLL-----EHRKLPK	113
Db	358	KEKQHIETQLLAERDLE-----RAEVAKATSHVGEIEQELALARDGHDQHVLEU--	406
Qy	114	NPQNTNPEGTSTGTPDEAAAPLAERTPDATABSLLKTPDEAARLVEGTSAKTQCLGKATSG	173
Db	407	-----EAKWDQLRTVVEAADREKVELLNQL--EEBKRVEDLQFRVE--ESITKG	453
Qy	174	KFEGSTETPRKILRPTKSEKFSWPAKERSKRIWTBEEKTSVKTECVAGVTNNKTVL	233
Db	454	DLETQTKLEHARI----KELEQSLLF--EKTAKDLQRELEDRVAT-----VSEKSRIM	502
Qy	234	EKGTSNNIACPTKET-----STKASTNVD--VSSVEPIFSLFG----TRTIENSQCT	279
Db	503	E--LEKDALURVQVAEIRRRLESNKPGVDMSLSLLQELISSLOEKLEVTRTDHQREIT	560
Qy	280	KVEEDFNLATKIISKSAQNYTCLPDATYQKDIK---TINHKIEDQMPSPSKREDEBY	336
Db	561	SLKEHFGAR-----EETHQKEIKALYATATEKLKENESLASKLEHANKE	604
Qy	337	SWDGSGLFESAKTQVCIPESMYQKME-----INREVELPE-KPAFPAVEMQ	386
Db	605	NSDVIALWKSLETAI----ASHQQAAMEELKVSPSKGLGTETABFAELKTQIERMLDYQ	660
Qy	387	KTVPN-----KAFELKNEQTLRAQMFPSESKQDKDEENSNDSESPCETVSKQVYL	438
Db	661	HEIENLQNDQDSRAAHAKMEALRAKLM-----KVIKEKENS-----LEAIRSK---L	706

Qy 439 PKA--THQKEDTTLGKLESPVND---GLLKPTCGRKVS-LPNKALELKDRETFKAESP 492
 Db 707 DKAEDQHLVEMEDTLNKLEAEIKVKELEVLQAKCNEQTVIDNFTSLQKATE-----759
 Qy 493 KDGGLKPTCGRKVSLPNKALELKDRETLK-----AESPDNDGLLKPTCGR 538
 Db 760 --EKLLDLDALRKASSSEKMKLRQQLQAEAEKIQIKHLEIEKNAESKASSITRELOGR 817
 Qy 539 KVSIPNKALELKD---RETEKAAQMFPSKQKDEENSWDFSEFLETLQNDVCLPKA 594
 Db 818 ELKLTNLQENLSEVQKTELEKELQILKFAEASEAVSQVSMQETV-----NK 869
 Qy 595 THQKE--PDTLSGKLESPDQGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVSTF 652
 Db 870 LHQKEQFNMLSDLEK-----LRENADMEAKFREK--DERBEOLIKAKE--KLNDI 919
 Qy 653 SLFGKPTTEN--SOSTKVEEDFNLTKEGATKTVTGQOERDI-----GIER 697
 Db 920 AEIMKMGDNGSOLTKMDELRL-----KERDVEELQLKLTKANENASFLQK 966
 Qy 698 APQDQTNKMTS-----ELGRKEDTKSTDSSEIISVSOTQNYECLP-BATYOK---E 745
 Db 967 SIEDMTVRAEQSQOAEAKKHEEKELERKLSLE--KKMETSHNQCOELKARYATSE 1024
 Qy 746 IKTTNGKLEESPEK-----PSHFEPATEMNSVNPKNGLKWKNTOTLRADSTTLKILDAL 800
 Db 1025 TKTRHEILQNLQKTLTDTEDKLKGARE-ENSGLLQLELEELRKQADKAKAQTAE--DAM 1081
 Qy 801 PSCERGRLLKKNCEQITAKMBQMKNFQVLQKELSEAKE--IKSQLENQAKWEQELCS 858
 Db 1082 QIME---QMTKEKETL-ASLEDTKTQNAKLQNELDTLKENNLKNVEELNKS--ELLT 1134
 Qy 859 VRLPLNQE-BEKRNVNLIK-----EKIRPEQLRKK-----889
 Db 1135 VE---NQMEEPREKIEITLQAAQKSOQLSALOENVKLABELGRSDEVTSHQKLEEE 1191
 Qy 890 -----LEVHGHOLEOTLRIODIELKSVTSNLNVSHTSENDLFHENCMLKKEIAML 941
 Db 1192 RVLNQLLEMKRRESKFIKDADEKASLOKIS-ITSALLTEKDAELEK-LRNEVTVL 1248
 Qy 942 KLEVALTKHQHOKENKYPEDIKI-LOEKNAELQMTLKLQ 981
 Db 1249 RGENASAKLSHVSQVOTLESQKLELKVKNLEQLKKNKQ 1289
 RESULT 2
 S67593
 N:Alternate names: protein D2552; protein YDL058w
 C:Species: Saccharomyces cerevisiae
 C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
 C:Accession: S67593; A38455; S30782
 R:Blöcker, H.; Brandt, P.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67587
 A:Accession: S67593
 A:Molecule type: DNA
 A:Residues: 1-1790 <BL0>
 A:Cross-references: UNIPROT:Q07380; EMBL:Z74106; NID:G1431058; PID:e253003; PID:G1431059
 A:Experimental source: strain S288C
 R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
 J. Cell Biol. 113, 245-260, 1991
 A:Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transp
 A:Reference number: A38455; MUID:91185402; PMID:2010462
 A:Accession: A38455
 A:Molecule type: DNA
 A:Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>
 A:Cross-references: GB:X54378; NID:G4777; PIDN:CAA38253.1; PID:G4778
 A:Note: the authors translated the codon ACT for residue 768 as Ile
 R:Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
 submitted to the EMBL Data Library, February 1993
 A:Description: An integrin analogue in Saccharomyces cerevisiae.
 A:Reference number: S30782

A:Accession: S30782
 A:Molecule type: DNA
 A:Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580, 'S'
 A:Cross-references: EMBL:L03188
 C:Genetics:
 A:Gene: SGD:USO1; INT1
 A:Cross-references: SGD:S0002216; MIPS:YDL058w
 A:Map position: 4L
 A:Keywords: coiled coil; transmembrane protein
 F:326-342/Domain: transmembrane #status predicted <TM1>
 F:394-410/Domain: transmembrane #status predicted <TM2>
 F:617-633/Domain: transmembrane #status predicted <TM3>
 Query Match 6.6%; Score 340.5; DB 2; Length 1790;
 Best Local Similarity 20.7%; Pred. No. 3.6e-07;
 Matches 250; Conservative 185; Mismatches 412; Indels 359; Gaps 53;
 Qy 1 MVAATLSYGAVI--EVQNKASITPLL-LAIQKRSQT-----VFPL---ITKANAN 46
 Db 601 VIKSLISFSYQIQEDVDVTKLVTLGLVAYEFSKESPPRKEFEFTIKTLGKNYAS 660
 Qy 47 AFNRSCTALMLAICESSSEIVGMLLQON-----VDVFAEDIHGITAERVAAR 95
 Db 661 RIQKFQKDSYFSKVDNEDSILPELDETGLPKYFSTVFIOLEFNIYRI---RTALSH 717
 Qy 96 GVNYY--IHQQLLEHIRKLPKNQNTNPEGTSTGTPDEAALPABERTPDTAESLLEK---TP 150
 Db 718 DPDEEPINKISPEVEKQLR--QCTKLKG-----EITSLOTETESTHENLTKLIALT 768
 Qy 151 DEARLVGTSKAIQICLGKATSKGFQO--STEETPRK-----ILRPTKSTSEKFSW 199
 Db 769 NEHKELDE---KQIILNSSHSLKENFSILELKNVRLSDLEMTQLRDVLQKQENQ 824
 Qy 200 PAKERSKITWEKETSVKTECVAGVTPNKTVELEKGTSNMIACPTKETSTKASTNVDS 259
 Db 825 TAL-LEKSTIHKQEDSIKT-----LEKLETLIS-----QKKAEADGINKM 865
 Qy 260 SVEPIFLSGTRTIENSQCTKVEEDFNLA-----TKIISKSAQAQNYTCLPDATYQDI 312
 Db 866 G-KDLPALSRQMAVEENCKNLQKEDKSNVNHQKETSLEKIDAAKIT-----EI 915
 Qy 313 KTIHNIKDQMFPSKREDEEYSDSGSILFESSAKTQVCIPESMYQKWEINREVEEL 372
 Db 916 KAINENLEEMKIQCNLSKKEHIS-----KEL 943
 Qy 373 PEKPSAFK-----PAVEMQKTVPN--KAFELKNEQTLRAAQMPFSEK-QKDDENSW 422
 Db 944 VEYKSRFQSHDNLVAKLTKLSLANNYKDMQAESNLKAVESKNESLQSLNQNKI 1003
 Qy 423 DSESPCTVTSQKVYLKPAHQKFDPLSGKLESPVKDGLLKPTCGRKVSLPNKALELK 482
 Db 1004 DS-----MSQE-----KENFQIBERSIEKNIEQ-----LKKT-----IS 1032
 Qy 483 DRETFKAESPDKGLLKPTCGRKVSLPNKALE-----LKDRETLKAESP 526
 Db 1033 DLEQTKEEIISKSSSDKDEYESQISLKEKLETATTANDENVNKISELTKEEELAEAL 1092
 Qy 527 DNDGLLKPTCGRKVSLPNKAL-ELKDRETFKAAQMPFSESKQKDDENSWDFSEFLETL 585
 Db 1093 AVKN-LKNELETKLETSEKALKVEKENBEHLKBEKIQLKEKATETKQOLNSLRANLESL- 1150
 Qy 586 QNDVCLPKATHQKFDPLTSGKLESPDKDGLLKPTCGMKISLPNKALELKDRETFKARDV 645
 Db 1151 -----EKHEHDLAQQLKYYE-----QIANKEQYNEEI 1179
 Qy 646 SSVSTFSLFGKPTTENSQSTKVEEDFNLTKEGATKTVTGQOERDIGIIEPAPQDQNK 705
 Db 1180 SOLNDEIT-----STQOENESIKKND-----ELEGEVKAAMKSTSEQSN-LKKSEIDALN- 1229
 Qy 706 MPTSELGRKEDTKSTG-----DSEIISVSDTQNYECLPEATYQKEIKTINGKIESP 757
 Db 1230 LQIKELKKKNETNEASLLSESIVSESTVKIKELQD-EC---NFKKEVSELEDKLASE 1285

Qy 758 EKPSHPFAPTEMQNSVFNKGLWKNKOTLRADSTTSLKILDALPSCRGRE-----L 809
Db 1286 DKNSKY---LELQKE-----SEKIBELDAKTTELKIQLEKITNLKAKESSELSRL 1336
Qy 810 KKNCEQ---ITAKMEQMKKFCV-----LOKELSE----- 837
Db 1337 KTSSEERKNAEQLKKEIKNEIQKQAFKPKLLNEGSSITTOEYSEKINTLEDELIR 1396
Qy 838 -----AKEI---KSQLE-----NOKAKWEQLCSVRLPLNQBEEK--- 869
Db 1397 LQENELKAKEIDNTRSELEKVSLSNDELLEEKQNTIKSLQDEILSYKDKITRNDKLLS 1456
Qy 870 -----RRNVILKEKIRPEB-----QURKKLVKHQLEQTLRI 902
Db 1457 IERNKRDLSLKBQLRAAQESKAKVBEGLKLBEESSKEKAELEKSKEMMKKLESTIES 1516
Qy 903 QDTELKSVTSLNQVSH-----THESENDLPHENCMKKEITAMLKLEVATLKHQHGV 954
Db 1517 NETELKSSMETIRKSDKELQSKKSAEEDIKNLOHESDLSIRINESEKIDIEELKSLRI 1576
Qy 955 --KENKYFEDIK---ILQEK---NAELQMTLKLKQKTVTKRASQYREQKVLTAENTML 1005
Db 1577 EAKSGSELETVKQELNNAQEKIRINAEENTVLKSLIEDIERELKDKQAEIKSNQOEKELL 1636
Qy 1006 TSKLKE 1011
Db 1637 TSLRKE 1642

RESULT 3
A43336
microtubule-vesicle linker CLIP-170 - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C:Accession: A43336
R;Pierre, P.; Scheel, J.; Rickard, J.B.; Kreis, T.E.
Cell 70, 887-900, 1992
A:Title: CLIP-170 links endocytic vesicles to microtubules.
A:Reference number: A43336; MUID:92405160; PMID:1356075
A:Accession: A43336
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1392 <PIE>
A:Cross-references: GB:M97501; NID:g180621; PID:AAA35693.1; PID:g180622

Query Match 6.58; Score 337; DB 2; Length 1392;
Best Local Similarity 22.28; Pred. No. 3.8e-07;
Matches 246; Conservative 181; Mismatches 401; Indels 280; Gaps 52;

Qy 1 MVATLLSYGAVIEQNKASLTPLLLAIQKRSKOTVEFLTKNANAFNESCCTALMLAI 60
Db 300 MATTSASLKRSPSASSLSMSSVASSVSSRPSRT--GLLTETSSRYARKISGTTALQAL 357
Qy 61 CEGSSEIVGMLLQNVDFAEIDHIGITAEYAAAGVNYHQHLLHIRKLPRNPQNTNP 120
Db 358 KERQOHEQLLAERDL-----ERAEVAKATSHV----- 385
Qy 121 EGTSTGTGPDNAAPLAERTPDTAESLLEKTPDEAARLVEGTS-AKIQCLGKATSGKFGST 179
Db 386 -----GEIQELALARDGHQHVLEBAKMDQLFTWVEAADREKVELLNQ-----L 431
Qy 180 EETPRKI---LRPTKETSEKFSW---PAKERSRKITWEEKTSVKTECVAGVTPNKTEV 232
Db 432 EEEKRVKVEDLQFRVEESITKGDLEVATVSEKSR-IMELEKDLALRVQEVA-----EL 483
Qy 233 LEKGTSMNIAPCTKETSTKASTNVVDVSSVBPFIPLFGTRTIENSQCCTKVEEDFNATKII 292
Db 484 RRRLESNK---PAGDWMSLSLLOEISSLOE--KLEVTRTDHOREITSKEHFGAR---- 534
Qy 293 SKSNAQNYTCLPDATYOKDIK---TINHKTEDQMPFSPESKKEEEDSVSDSGSLFESSAK 349
Db 535 -----EETHQKEIKALYATATEKLSKENESLKSLEHANKENSVDIALWKSLE 582

Qy 350 TOVCIPESMYQKQVME-----INREVELPE-KPSAFKPAVEMQKTVPN----- 391
Db 583 TAI-----ASHQQAAMEELKVSFSGKGLGTETAFABELKQTQIEKWRLDYQHEIENLQOQDSE 638
Qy 392 KAPELKNEQTLRAAQMPFSPESKQKQDENSWDSSEPCETVSKQDVYLPKA--THQKEPDT 449
Db 639 RAAHAKMEALRAKLM-----KVTKERENS-----LEAIRSK---LDKABDQHLVEMED 684
Qy 450 LSGKLBESPVKD---GLLKPTCGRKVS-LPNKALELKDRETFKAEAPDKDGLLKPTCGRK 505
Db 685 TLNKLQAEAIKVKLEVLQAKCNEQTKVIDNFTSOLKATE-----EKLDDLALRK 735
Qy 506 VSLPNKALELKDRTELK-----AESPDNDGLLKPTCGRKVSIPNKALEKLD 551
Db 736 ASSEKSEMKKLQOLQAEABKQIKHLEIEKNAESSKASSITRELQGRRELKLTNLQENLSE 795
Qy 552 ----RETFKAAQMPFSPESKQKQDENSWDFESFLETLLONDVCLPKATHOKE--PDTLSG 605
Db 796 VSQVKETLEKQLTLKEKFAEASEEAVSVQKSMQETV-----NKLHQKEBQFNMLUSS 847
Qy 606 KLBPSPDKDGLLKPTCGMKISLPNKALELKDRETFKAEADVSSVESTFSFGKPTTEN-SQ 664
Db 848 DLEK-----LRENLMADMEAKFREK--DEREEQLIKAKE--KLENDIAEIMKMSGDSNSQ 897
Qy 665 STKVEEDFNLTKEGATKTVTGOQERDI-----GIIRAPQDQTNMPTSS- 709
Db 898 LTKMNDLRL-----KERDVEELQLKLTKANENASFLQKSIEDMTVKAESQS 944
Qy 710 -----ELGRKEDTKSTDSEIISVSDTQNYECLP-ENTYQK---ELKTTNGKIEESPE 758
Db 945 QEAAKGEHEEKKELERKLSDLK--KQMETSHNOCEQLKARYATSETTKTTHSEILQNLQ 1002
Qy 759 K-----PSHFEPATEMQNSVFNKGLWKNKOTLRADSTTSLKILDALPSCRGRELKKN 813
Db 1003 KTLTDTEDKLKGARE-ENSGLLQLEBLRQAEAKAAQTAE--DAMQIME---QMTKEK 1056
Qy 814 CEQITAKMEQMKKFCVLQKELSEAKE--IKSLENOKAKWEQBELCSVRLPLNQB-BEKR 870
Db 1057 TETL-ASLEDTKQTNAKLQNELDTLKENNLKNVVELNKS--ELLTVE---NQMESFR 1109
Qy 871 RNVDLIK-----EKIRPEEQLRKK-----LEVKH 894
Db 1110 KEIETLKQAAAKSQSLQALQENVKLAELGSRDEVTSHQKLEEBERSVLNQLLEMKK 1169
Qy 895 QLEOTLRIDQIELKSVTSLNNOVSHSTHESNDLPHENCMKKEITAMLKLEVATLKHQHGV 954
Db 1170 RESKFIKDADEEKASLOKSIS-ITSALLTEKDAELEK--LRENVTVLGENASAKSLHSV 1226
Qy 955 KENKYFEDIKI-LOEKNAELQMTLKIKQ 981
Db 1227 VQTLSEDKVKLELKVKNLELQLENKRQ 1254

RESULT 4
JC5421
smooth muscle myosin heavy chain 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C:Accession: JC5421
R;Hasegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A:Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A:Reference number: JC5420; MUID:97242182; PMID:9125171
A:Accession: JC5421
A:Molecule type: mRNA
A:Residues: 1-1938 <HAS>
A:Cross-references: UNIPROT:O08638; DDBJ:D85924; NID:g1945079; PID:BAAL19691.1; PID:g194
A:Experimental source: smooth muscle
C:Comment: This protein plays a role in smooth muscle cell contraction.
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: nucleotide binding; P-loop
F;88-771/Domain: myosin motor domain homology <MMOT>
F;178-185/Region: nucleotide-binding motif A (P-loop)

Query Match		6.2%; Score 321.5; DB 2; Length 1938;
Best Local Similarity		20.7%; Pred. No. 2.6e-06;
Matches		232; Conservative 184; Mismatches 438; Indels 269; Gaps 42;
Qy	73	QONVDVFAEDIHGTAERYAAARGVNI---HQOLLEHIRKLPKNPONTNPEGTSTGTPD 129
Db	851	EEENQAKEEMQKITERQQAETELKEQKHTQLAEKTLLOEQLO-----AETELYA 904
Qy	130	EAAFLAERTPTDAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQ-----STE 180
Db	905	ESEEMRVLAACKQELBEILHEMEARLEEDRQQL--QAERKMAQQMLDLEEQLEEE 962
Qy	181	ETPKILRPTKETSSEKFSWPAK-----ERSKLTWEEKETSVKTECVAGVTN 228
Db	963	EAAQKQLQLEKVTAE-----AKIKLEDDILVMDQNSKLS---KERKLEERVSDITN 1014
Qy	229	KTEVLEKGTSMIACPTKETSTKASTNV-----DVSSV-EPIFS 266
Db	1015	LAEBEEK-AKNLTKLKSHESMISELVRLKKEKSQELKRLKLEGDASDPHEQIAD 1073
Qy	267	LFGRRTIENSQCTKVEEDFNLATKIISKSAQNTYCLPDATYQKDITKINHIEDQMPFS 326
Db	1074	LOAQIAELKQLAKKEELQAALARLDEBETIAQKNAL-----KKIRELEGHISDLQEDL 1127
Qy	327	ESK---REDEEYSDWSGLSPESAKTQV--CIFESMYQKWEINRE-----368
Db	1128	DSERAARKAEKQKRDGLGEEL-ALKTELEDTLDTSTATQBELRAKRGQEVTVLKKALDEE 1186
Qy	369	-----VEELPEKPSAFKPAV-----EMQKTPNKAPELKNQOTLRAAQ 406
Db	1187	TRSHEAQVOEMRQKHTQAVBELTQELQFKAKANLDSKQTKLEKANADLAGELRVLGOA 1246
Qy	407	MPPSESKQKDDENSWDSESPCTVSKQDVLPKATH--OKEFDTLSKLEESPVKQGLL 464
Db	1247	KQEVHEHKKKLEVLQDLQSKCSDGERARAEISDKVHKLQNEVESVTGMLNEA-----1299
Qy	465	KPTCGKRVSLPNKALEL-KDRETFKAESPDKDGLLKPTCGKRVSLPNKALELKDRITLKA 523
Db	1300	-----EGKATKLVADVASGLQDTQELLOEETROKLVSTKRLQLEDERNSLQ 1349
Qy	524	ESPNDGLLKPTCGKRVSLPNKALELKDRITFFKAAQMPFSESQKDDENSWDFESFLET 583
Db	1350	DQDDEMEAKQNLERHVSITLNIQL-----SDSKK-----LQDFASTTIEV 1389
Qy	584	LLQNDVCLPKATHQKEFTLSKLEE-----SPDKDGLLKPTCGMKISLPNKA-----631
Db	1390	MEEG-----KKRLQKEMEGLSQOYBEKAAAYDKLEKTKNRLQOELDDLVDLDNQRLVS 1444
Qy	632	-LELQDR-----ETPKAEDVSVSEFTSLFGKPTTENSOSTKVEEDF 672
Db	1445	NLEKQKQKPDLLABEKNTSSKYADRRDAEABAREKETALSARALE--EALAEKSEL 1502
Qy	673	NLTTKK-----GATKVTVQOERDIGIIBRAPQDQTNKMPSTSELGRKEDTKTSDSE 724
Db	1503	ERTNMLKAMEDLVSSKDDVGNVHLEKSKRALETQMEEMKTOLEESDDVQATEDAK 1562
Qy	725	I-ISVSDTONYECLPATYQKIKITNGKLESP---EKPSHFEPATQNSVPNKGLEW 780
Db	1563	LRLEV-----NMQAL-KGQFERDLOARDEQNEEKRRQLQRLH-EVETELEDERKQALAA 1616
Qy	781	KNKOTLRADSTTLKDALPSCGERELKDNCEQITAKMEQMKNKFCVLQKELSEAKE 840
Db	1617	AKKKLEGLDKOLEAD---SAIKGRE---BAIKQLRKLQAKMD---FQRELDARA 1666
Qy	841	IKSQL-----ENQK--AKWEQELCSVRLPINQEEKRNVDILKEKIRPE-----883
Db	1667	SRDEIPATSKENEKKAKSLEADLMQLEDLAAERAKQADLEKEELAEELASSLGRNT 1726
Qy	884	-EQRLKLELVK-HQLQTLRIQDIELKS-----TSNLQNVSHTHESENDLFHEN---931
Db	1727	LODEKRLREARIAQLBEEEEEEOGNMEASDRVRKATLQAEQSNELATERSTAQNESA 1786

Qy	932	-----CMLKKEIAMLKLEVALTKHQHVQKENVKENVKEDIKILOE 968
Db	1787	ROOLERQNKELSKLOEVEGAVKAKLKSTVALEAKIAQLEQVEQEAEREKQAATKSLKQ 1846
Qy	969	KNAELQMTLKLKQKVTTKRASQYREQLKVLTAENTMLTSLKLE 1011
Db	1847	KDKKLKEVL-LQVEDERKMAEQYKEQAEGKNTKVQLKQLEEBE 1888
RESULT 5		
JCS420		
smooth muscle myosin heavy chain 1 - mouse		
C;Species: Mus musculus (house mouse)		
C;Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004		
C;Accession: JCS420		
R;Hasagawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.		
Biochem. Biophys. Res. Commun. 232, 313-316, 1997		
A;Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.		
A;Reference number: JCS420; MUID:97242182; PMID:9125171		
A;Accession: JCS420		
A;Molecule type: mRNA		
A;Residues: 1-1972 <HAS>		
A;Cross-references: UNIPROT:O08638; DDBJ:D85923; NID:g1945077; PIDN:BAA19690.1; PID:g194		
A;Experimental source: smooth muscle		
C;Comment: This protein plays a role in smooth muscle cell contraction.		
C;Superfamily: myosin heavy chain; myosin motor domain homology		
C;Keywords: nucleotide binding; P-loop		
F;88-771/Domain: myosin motor domain homology <MMOT>		
F;178-185/Region: nucleotide-binding motif A (P-loop)		
Query Match		6.2%; Score 321.5; DB 2; Length 1972;
Best Local Similarity		20.7%; Pred. No. 2.7e-06;
Matches		232; Conservative 184; Mismatches 438; Indels 269; Gaps 42;
Qy	73	QONVDVFAEDIHGTAERYAAARGVNI---HQOLLEHIRKLPKNPONTNPEGTSTGTPD 129
Db	851	EEENQAKEEMQKITERQQAETELKEQKHTQLAEKTLLOEQLO-----AETELYA 904
Qy	130	EAAFLAERTPTDAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQ-----STE 180
Db	905	ESEEMRVLAACKQELBEILHEMEARLEEDRQQL--QAERKMAQQMLDLEEQLEEE 962
Qy	181	ETPKILRPTKETSSEKFSWPAK-----ERSKLTWEEKETSVKTECVAGVTN 228
Db	963	EAAQKQLQLEKVTAE-----AKIKLEDDILVMDQNSKLS---KERKLEERVSDITN 1014
Qy	229	KTEVLEKGTSMIACPTKETSTKASTNV-----DVSSV-EPIFS 266
Db	1015	LAEBEEK-AKNLTKLKSHESMISELVRLKKEKSQELKRLKLEGDASDPHEQIAD 1073
Qy	267	LFGRRTIENSQCTKVEEDFNLATKIISKSAQNTYCLPDATYQKDITKINHIEDQMPFS 326
Db	1074	LOAQIAELKQLAKKEELQAALARLDEBETIAQKNAL-----KKIRELEGHISDLQEDL 1127
Qy	327	ESK---REDEEYSDWSGLSPESAKTQV--CIFESMYQKWEINRE-----368
Db	1128	DSERAARKAEKQKRDGLGEEL-ALKTELEDTLDTSTATQBELRAKRGQEVTVLKKALDEE 1186
Qy	369	-----VEELPEKPSAFKPAV-----EMQKTPNKAPELKNQOTLRAAQ 406
Db	1187	TRSHEAQVOEMRQKHTQAVBELTQELQFKAKANLDSKQTKLEKANADLAGELRVLGOA 1246
Qy	407	MPPSESKQKDDENSWDSESPCTVSKQDVLPKATH--OKEFDTLSKLEESPVKQGLL 464
Db	1247	KQEVHEHKKKLEVLQDLQSKCSDGERARAEISDKVHKLQNEVESVTGMLNEA-----1299
Qy	465	KPTCGKRVSLPNKALEL-KDRETFKAESPDKDGLLKPTCGKRVSLPNKALELKDRITLKA 523
Db	1300	-----EGKATKLVADVASGLQDTQELLOEETROKLVSTKRLQLEDERNSLQ 1349
Qy	524	ESPNDGLLKPTCGKRVSLPNKALELKDRITFFKAAQMPFSESQKDDENSWDFESFLET 583
Db	1350	DQDDEMEAKQNLERHVSITLNIQL-----SDSKK-----LQDFASTTIEV 1389

A26655 myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N:Contains: myosin Atpase [EC 3.6.4.1]
C:Species: Dictyostelium discoideum
C>Date: 05-Oct-1988 #sequence revision 05-Oct-1988 #text_change 09-Jul-2004
C:Accession: A26655; A24728; S00250
R:Warrick, H.M.; De Lozanne, A.; Lewin, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A:Reference number: A26655; MUID:87092266; PMID:3540939
A:Accession: A26655
A:Molecule type: DNA
A:Residues: 1-2116 <NAR>
A:Cross-references: UNIPROT:P08799; GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1;
R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Lewin, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A:Reference number: A24728; MUID:86016788; PMID:3901008
A:Accession: A24728
A:Molecule type: DNA
A:Residues: 2035-2116
R:Waggle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A:Reference number: S00250; MUID:88112226; PMID:2828113
A:Accession: S00250
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1734-1893 <WAG>
C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphate binding site
F:1-818/Domain: myosin motor domain homology <WAG>
F:89-747/Region: actin binding site
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:819-2116/Domain: alpha-helical rod <ROD>

Query Match 6.2%; Score 318.5; DB 2; Length 2116;
Best Local Similarity 20.4%; Pred. No. 3.9e-06;
Matches 210; Conservative 181; Mismatches 376; Indels 263; Gaps 44;

QY 73 QNVVDVFAEDIHGTAERYAARGVNYHQHLEHRL-----PNQNTNPEGTS--T 125
DB 1201 KKKVELDLEKSAQLAEATAQAQALDKKKLEQELSEVOTQLSEANNKVNDSNTKHL 1260
QY 126 GTPDEAALPAPTPTDAELLEKTPDEAARLVGTSAKIQCGLKATSGKFGESTETPRK 185
DB 1261 ETSFNKLKLEAEQAKQALEK-----KRL--GLESELKHVN-----EQLEE--- 1302
QY 186 ILRPTKETSEKFSWPAKERSKRITWBEKETSVKTECVAGVTENK--TEVLEKGTSNMIAC 243
DB 1303 --KKQESNEK-----RKVDLEKEVSELKQDIEEVASKAVTEAKNKKSELDDEI 1351
QY 244 PTKETSTKASTNVDSVSPISFLFTRTIENSQCTKVEEDFNATKIISKSAANYTCL 303
DB 1352 KRQYADVSSRD--KSVEQLTKLQAKN--EELNRTAEAEQGLDRAERSKKKA----- 1400
QY 304 PDATYQKDIKTHNKLELDQMFPESEKREDEEYSDWSGLFESSAKTQVCIPESMYQKVM 363
DB 1401 -EFDLEAEVKNLEETAKKAKKAKKAAETDTRTSKSEL--DDAKN--VSEQYQVQIK 1454
QY 364 EINREVEE---LPEKPSAFKPAVEMQKTPNKAFLKNEQTLRAAQMFPSSKQKQDE- 418
DB 1455 RLNEELSELRSVLEADERCNSAIKAKTAESALESILKDB--IDAANNAKAAERKSKEL 1512
QY 419 -----ENWDSSES---PQETVSQKDVYLPKATHQKEPDTLSGKLE---ESPVKDGLLK 465
DB 1513 EVRVAELESLEKSDSGTWNVEFRKKDA-----EIDDLRARLDRETESRIKD--- 1560
QY 466 PTCGRKVSFLPKALEKDRFTFAESPDQDGLLKPTCGRKVSFLPNKALELKDRITLKAES 525
DB 1561 -----EDKKNRKQFPADLEA-----KVREAQREVVITDRKKKLES 1596
QY 526 PDND---GLLKPTCGR-KVSLPNKALE--LKDRFTFAEQMFPSESKQKQDDENSWDFES 579

DB 1597 DIIDLSTQDITETSKRIKIEKSKKLEQTLAER---RAAEEGSKAADEEIRKQWQEV 1653
QY 580 FLETLLQNDVCLPKATHOK-----EPDTLSGKLE-ESPDQDGLLKPTCGMKISLPNKAL 632
DB 1654 ELRAQLDSEERAAALNASEKKIKSLVAEVDVEKQLEDEILAKDKLVKAKRALEVELEVRD 1713
QY 633 ELKDRFTFAE-----DVSSVESTFSLFGKPTTENSQSTKVEEDFNLTTKEGATK 682
DB 1714 QLEEEEDSSELEDSKRLTTEVEDIKKKYD-----AEVEQNTKLDE-----ARK 1758
QY 683 TVTGOQERDIGIIRAPQDQTNMPTSELGRK-----EDTKSTSDSEIISVSDTQNYEC 736
DB 1759 KLTD-----DVTDLKKQLEDEKKKLMESEKRAKRLSENEDEFLAKLDAEVKNRSRAEK--- 1811
QY 737 LPEATYQKEIKTTNGKIBESPEKSHFEPATEMQNSVPNKGLEWKNKQTLRADSTLSKI 796
DB 1812 -DRKKYEKDLKDTKYKIND-----EAATQTEIGAAKLE----- 1845
QY 797 LDALPSCERBELKKNCEQITAKMEQMKNCVLOKELSEAKEIKSLENOKAKWEQEL 856
DB 1846 -----DQIDELSKLEQEQAKATQADK---SKTLEGEIDNLRQAQIEDS- 1886
QY 857 CSVRLPLNQBEERK---NVDILKEKIRPEEQLRKKLEVKHOLEOTLRIQDIELKSVTSN 913
DB 1887 --GKIKVRLEKRALEGELEELRETVEAE-----DSKSEAEQSKRLVELELEDARN 1938
QY 914 LNQVSHTHESENDPLHENCMLKKEIA-----MLKLEVATLKHQ- 951
DB 1939 LQKEIDAKEIAED---AKSNLQREIVAEKGRLEESIARTNSDRSRKRLAEAITDALTAQV 1995
QY 952 -----HQVKENKYFE-DIKILOKRNALQMTLKLKOKTYTKRASQYR-----BQ 994
DB 1996 DAEQAKANQOIKENKIKETELKEYRKFGSEKTKTKEFLVWEKLETDYKRAKKEADEQ 2055
QY 995 LKVLTAENTM 1004
DB 2056 QORLTVENDL 2065

RESULT 8
A41604 myosin heavy chain, smooth muscle, long splice form - rabbit
N:Contains: myosin Atpase [EC 3.6.4.1]
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A41604; A33501
R:Babij, P.; Kelly, C.; Periasamy, M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10676-10680, 1991
A:Title: Characterization of a mammalian smooth muscle myosin heavy-chain gene: complete
A:Reference number: A41604; MUID:92073350; PMID:1961735
A:Accession: A41604
A:Molecule type: mRNA
A:Residues: 1-1972 <BAB>
A:Cross-references: GB:M77812
R:Nagai, R.; Kuro-o, M.; Babij, P.; Periasamy, M.
J. Biol. Chem. 264, 9734-9737, 1989
A:Title: Identification of two types of smooth muscle myosin heavy chain isoforms by cdn
A:Reference number: A33501; MUID:89255535; PMID:2722872
A:Accession: A33501
A:Molecule type: mRNA
A:Residues: 1455-1972 <NAG>
A:Cross-references: GB:J04833; NID:g165519; PIDN:AAA31407.1; PID:g165520
A:Experimental source: smooth muscle
A>Note: examination by Southern blotting for the regions of difference between this isoform
active splicing
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated
F:88-771/Domain: myosin motor domain homology <WAG>
F:178-185/Region: nucleotide-binding motif A (P-loop)
F:559-572/Region: actin binding #status predicted
F:633-647/Region: actin binding #status predicted
F:844-1938/Domain: coiled coil #status predicted <COI>

F:1285-1972/Region: S2
F:1285-1972/Domain: light meromyosin
F:1285-1972/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted
F:129/Binding site: ATP (Lys) #status predicted
F:184/Binding site: ATP (Lys) #status predicted
F:701,711/Active site: Cys #status predicted

Query Match 6.1%; Score 315; DB 1; Length 1972;
Best Local Similarity 20.1%; Pred. No. 5.1e-06;
Matches 227; Conservative 203; Mismatches 46; Indels 242; Gaps 41;

Qy 13 EVQNKASITPLLLAIAKESKOTVE--ITAEYAAAGVNYIHQQLLEH-----IRKL 111
Db 866 ERQQAASE--LQELQKHTQLSEKQLQAEKQWAOQLDLEQLSEEAARQKLEKVTAAKIKKL 983
Qy 71 LLOQNVDFABDIHG--ITAEYAAAGVNYIHQQLLEH-----IRKL 111
Db 924 LHMEARLEEDRQQLQAEKQWAOQLDLEQLSEEAARQKLEKVTAAKIKKL 983
Qy 112 P-----KNPQNTPEGSTGTPDPAAPLAERTPTASLLEKTPDEARLVEGTSKIQIC 166
Db 984 EDDILVMDQNNK-----LSKERKLEERISDLTTLNLAEE--BEKAKNLTK----- 1027
Qy 167 LGKATSGKEFQSTETPRKILPTKETSEKSWAKERSRKITWEEKSVK-TECVAGV 225
Db 1028 ----LKNHESNISELVRLKKEKSRQE-----LEKLKRWGGEASDLHEQIADLQAOI 1078
Qy 226 TPNKTEVLEKGTSMNIACPTKETSTKASTNVDSVSEPFISLFGTRTIENSQCTKVEEDF 285
Db 1079 AELKQWLAKKEELQALARLEDETSQKN-----ALKKIRELE-CHISDLQEDL 1127
Qy 286 NLATKIISKSAQNYTCLPDATYQKDIKTINHIEDQMFPSSEKRE-----EDEYSWDSG 341
Db 1128 D-----SERAARKAEKQKRDIGEELEALKTELEDLTDTATQQLRAKREQEVTVLKK 1181
Qy 342 SLFESSAKTQVCIPESMYQKWEINREVEELPEKPSAFKPAV-----EMQKTPNKAFLK 397
Db 1182 ALDEBTRSHQAQOE--MRQKHTQV---VEELTEQLEQFKRANKANLDTKOTLEKENADLA 1237
Qy 398 NEQTLRAQMPSPSSKOKDDENSWDSSEPCETVSQKDVLPKATH--QKEFDTLGKLE 455
Db 1238 GELRVLGQAQKQEVHKKKLEVLQELQSKCSGERARAEALNDVKHKLQNEVESVTVGLS 1297
Qy 456 ESPVKGDLKPTCGRKVSLPNKALEL-KDRETFKAESPDKDGLKPTCGRKVSLPNKALE 514
Db 1298 EA-----EGKAIKAKEVASLGSQLDQTQELQEBETROKLVNSTKLRLQ 1340
Qy 515 LKDBETLKABSPNDGGLKPTCGRKVSLPNKALELKDRETFKAQMPSPSSKOKDDENS 574
Db 1341 LEDERNLQQLDDEMEAKQNLRIHISTLNLQ-----SDSKKK-----L 1380
Qy 575 WDPESFLETLNDVCLPKATHQEFDTLSGKLEE-SPDKDGLKPTCGMKISLPNKALE 633
Db 1381 QDFASTVESLEEG-----KKRQKEISLTQOYEKAAAYDKLEKTVNRQOELDLVD 1435
Qy 634 LKDBETKABDVSSVESTSLFGKPTTENSQ-STKVEED-----FNLITKEGATKVTQOQE 689
Db 1436 LDNRQL-----VSNLEKKQKFDQLAAEKNISSKYADERAEAREKETSALSLARA 1491
Qy 690 RDIGIIEAPODQNTKMPSTBELGRKEDTKSDSEIISVSDTQNYECLPEATYQKEIKT- 748
Db 1492 LEEALEAKEELERTNKLKABM-----EDLVSSKDDVGKNVHELEKSKALE--TQMEEMKTQ 1547
Qy 749 -----TNGKIEESPEKPSHF-----EPATEMON 771
Db 1548 LEELEDELOATEDAKLLEVNMQALKVQFERDLQARDEQNEKRRQIQORLHVEYTELED 1607
Qy 772 SVPNKGLWKNQOTLRADSTTLTKILDALPSCERGLKKNDCNCRQITAKMEQMKKFCVL 831
Db 1608 ERKQALAAAAKKLEGLDKDLELQAD---SAIKGRE---EAIQKLLKLAQAKMD----F 1657
Qy 832 QKELSEAKETKSQI-----ENQK---AKWEQLCSVRPLNQEEBEKRRNVDILKIRPE- 883

Db 1658 QRELEDAARDEIFATAKENEKAKSLEADLMQLOQEDLAAARARQADLEKELABEL 1717
Qy 884 -----EQLRKLEVK-HQLEQTLRIQDIELKSV-----TSNLNQVSHTHESN 925
Db 1718 ASSLSGRNALQDEKRRLEARIQLEEBEEOGMEASDRVRKATQOAEQLSNELATER 1777
Qy 926 DLPHEN-----CMLKKEIAMLKLEVATLKHQHVQKENVY 959
Db 1778 STAQKNESARQQLERQNKELSKLOENEGAVKSKFKSTIAALEAKIAQLESQVQEAAREK 1837
Qy 960 FEDIKILOENAELOMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011
Db 1838 QAAAKALQRDKKLKLEML-LQVEDERKQABQYKEQAEKGNKAKVKQLKQLEEE 1888

RESULT 9
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C/Accession: T34418
R/Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A/Description: The sequence of C. elegans cosmid F12F3.
A/Reference number: Z21521
A/Accession: T34418
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-3488 <FUL>
A/Cross-references: EMBL:U80022; PIDN:AAC25985.1; GSPDB:GN00023; CESP:F12F3.3
A/Experimental source: strain Bristol N2; clone F12F3
C/Genetics:
A/Gene: CESP:F12F3.3
A/Map position: 5
A/Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 6.1%; Score 315; DB 2; Length 3488;
Best Local Similarity 22.0%; Pred. No. 1e-05;
Matches 222; Conservative 159; Mismatches 423; Indels 203; Gaps 41;

Qy 120 PEGTSTCTPDEAPLAERTPTDAESLLEKTPDEAARLVEGTS-----KIOCLGKATSGKF 175
Db 511 PDEISSRFGPSTWHSETNITTTIRREGSADAKTPLVERPLSASVMKVFVTLVESAKEKA 570
Qy 176 EQS-----TEETP-----RKILRPTKETSEKSWPAKERSRKITWEEKSVKTECVAGVT 226
Db 571 EFSFKRSETPDDKSRKKEGLPPAKSEKKEDEVTAEQSTEALLESKKEVDESKISEQQ 630
Qy 227 P-----NKTVEL-----EKGTSMNIACPTKETSTKASTNVDSVSEPFISLFGTRT 272
Db 631 PSDKNKSEVVGVPEKAAGPETKKQVSEIEEVPVKKTIKKKTEKSD--SSISQKSNVLKPAD 689
Qy 273 IENS-----QCTKVEEDFNATKIISKSAQNYTCLPDATYQKDIKTINHIEDQMPF 325
Db 690 DDKSKSDVTVTKSKKTTED---QTKVATDSKLEK---AADTTKQIETETV---VDD----- 736
Qy 326 SESKREDEYSWDSGSLFESSAKTQVCI-----PESMYQKWEINREVEELPEKPSAFK 380
Db 737 -KSKKVLKKTTEKSDSFISQKSETPPVPEPTPAESEAQKIAEVNKAQKQEVDDNLKR 795
Qy 381 PAVEMQKTPNKAFLKNEQTLRAAQMPSPSSKQKQ--DEENSWDSE-----SPCETVS-Q 433
Db 796 EAEVAAKIADEKLKIEAEANIKKTAEEVAAKQKQEKDEQLKLETEVVVSKSAKLELE 855
Qy 434 KDVLPLKATH-----QKEFD-----TLGKLEESPVKDGLL 464
Db 856 KQAOIKKAAEADAVKQKELNKNKNEKBAKSAADKLEESAASKKSVSESVKFGEE 915
Qy 465 KPT-CGRKV-----SLP--NKALSLKDRTEFKA--ESPDKGLLKPTCGRKVSLPNKALEL 515
Db 916 KTKAGEKTVQVSEPTSKTIDTKDVGATEPADETPKKIIKKTKESDSSISQKS--A 973
Qy 516 KDBETLKABSPNDGGLKPTCGRKVSLPNKALELKORET---FKAAQMPSPSSKQKODEE 572

R; Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M. Mol. Cell. Biol. 14, 2564-2576, 1994

A; Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein

A; Reference number: A56539; MUID:94187728; PMID:7511208

A; Accession: A56539

A; Molecule type: mRNA

A; Residues: 1-3259 <SEE>

A; Cross-references: UNIPROT:Q14789; EMBL:X75304; NID:9405714; PIDN:CAA53052.1; PID:9405714

C; Genetics:

A; Gene: GDB:GOLGB1; GCP; GCP371

A; Cross-references: GDB:454958

A; Map position: 3q13.31-3q13.31

C; Superfamily: glanxin

C; Keywords: coiled coil; Golgi apparatus; transmembrane protein

F; 3238-3254/Domain: transmembrane #status predicted <TMN>

Query Match 6.0%; Score 311; DB 1; Length 3259;
Best Local Similarity 19.3%; Pred. No. 1.4e-05;
Matches 235; Conservative 212; Mismatches 425; Indels 346; Gaps 48;

QY 14 VQNKASLTPLLLAIQKRSKQTVBFLITKMANANAFNE---SKCTALMLAICEGSSIIIVGM 70
DB 1101 LQDKTQIDLLQAEISENQAIQKLITSNTDSDGDSVALVKEVTVISPPCTGSSHHKP 1160
QY 71 LLOQNV---DVFAEDIHGITAERYAAAGVNYIHQQLLEHIR---KLPRKNPQNTPEGTS 124
DB 1161 ELBEKILALEKEXEQQLKQLEALTSRKALKKAQAEKREHLREELKQKQDDYNRLQEQF- 1219
QY 125 TGTPTDEAAPLAERTPDAESLLSEKTPDEAARLVEGTSAKIQCLGKATSGKFQESTETPR 184
DB 1220 ----DEQSKENENIGQLRQLQIVRESIDGKLPSLPDQDESC--SSTPGLEELPKATEQ 1273
QY 185 KILRPKTESKESFPWAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTFS-NMIAC 243
DB 1274 HHTQPVLESNLCDFWDSHSD-----ASALQGTSTVAQIKA 1309
QY 244 PTKE-TSTKASTNVVDSSVEPIFSLFTRPTIENSQCTKVEEDFNLATKIISKSAQNYTC 302
DB 1310 QLKEIEAEKVELEKLVSS-----TSELTKKSEEVFQLQEQINKQGLE---- 1352
QY 303 LPDATYQKQIKTINHKE-----KTOVCIPESMY-----DOMFPSESKREE 332
DB 1353 -----TESLKTVSHAEVHAESLQOKLESQIQIAGLEHLRELQPKLDELQKLSKKEE 1406
QY 333 DEEYSDSGSLPSSA-----KTOVCIPESMY-----QKWEINRREVEL 372
DB 1407 DVSYL--SGQLSEKAALTKIQTEIIEQEDLIKALHTOLEMQAKEHDERIKQVQLCEM 1464
QY 373 PEKP-----SAPKPAVE--MOKTVPNKAPFELKNEQTLR-----AAQMPFSEKQKDDBE 419
DB 1465 KQKPEIGESRAKQIQIKLQAALISRKALKENKSLQBELSLARGTIERLTKSLADVE 1524
QY 420 NSWDSSEPCETVSKDYLPLK-ATHQKEPDTLSCKLEESPVKDGLLKPTCGRKVSLPNKA 478
DB 1525 SQVSAQN-----KEKOTVLGRLLAQEERDKLITEMDRSLLENQSLSSC-----ES 1571
QY 479 LEUKDRETFKAESPDKGLLKPTCGRKVSLPNKALELKOR-----518
DB 1572 LKL-----ALEGLTEDKELVETESLSSKIAESTEWQEKHKELOKQEYELLQSYENVSN 1627
QY 519 -----ETLKAESPNDGLLKPTCGRK-----VSLPN 544
DB 1628 EAERIQHVVFAVQEQELVGLKLRSTEANKKETEQQLQAEQEQEEMKEKWRKFAKSKQ 1687
QY 545 KALELKDRETFKAQAMPFSESKQKDDDENSWDPESFLETLQNDVCLPKATH--QKEFDT 602
DB 1688 KILELSEENDRLRAEVHPAGDTAKE-----CMETLLSSNAGMKBELERVKMEYET 1737
QY 603 LSKCLEE-SPDKDGLLKPTCGMKISLPN-----KALELKDRETFKAED-----VS 646
DB 1738 LSKKQSLMSKSEKDEEVODLKHQIEDNVNSQANLEATEKHDNQNTVNTVEEQTSIPGET 1797
QY 647 SVBESTFSLGKPT--TENSQSTK-----VBEDFNLTTK-EGATKTVTGQOERDQIGITERAP 699

```

Db      1798  EQDSLSMSTRPTCSSEVPSAKSANPAVSKDFSHDEINNYLQOIQDLKERIAGLE--- 1854
Qy      700  QODTNKMPSTSELGRKEDTK-----STSDSETISVSDTQNYECLPEATVQKEIKTNGKIEE 755
Db      1855  EKQKNKEFGOTLNEKNKTLISQISTQDGLKMLQEEVTKWNLNQQIQEELSRT-KLKE 1913
Qy      756  SPEKPSHFHPATEMQNSVPNKGLEWKNKQTLRADSTTLTKILDALPSCERGELKKNCE 815
Db      1914  TAE-----EKKDLEERLNNQ-----LAELNGSI-----GNYCQ 1942
Qy      816  QIT---AKNQMKNCFCVLQKELSEAKEIKSQLENOKAKWEQELCSVRL-----PL 863
Db      1943  DVTDAQIKNELLESEMKNLKCVCSELEERKQQLVKEKTKVESEIRKEYLEKIQGAQKEPG 2002
Qy      864  NOEEERKRVNVDILKEKIRPEEQRLKKLVKQK-----LEOTLRIQDI-----ELKSVT 911
Db      2003  NKSHAKELQ-ELKKEQKQEVKQKQKDC-IRYQEKISALERTVVALEFPVQTESQKDLKITK 2060
Qy      912  SNLNQ-VSHTHSENDL-----PHENCMLKKEI-----AMLKLEV 945
Db      2061  ENLAQAVEHRKKAQAELASFVLLDDTQSEARVLADNLKLKKELOSNEKSVKSMQKQD 2120
Qy      946  ATLKHQHQVKENKYFDDIKILQEK-----NAELQMTLKLKQKTVTKRASQY 991
Db      2121  EDLERRLEQAEEKHLKEKNMQEKLDAALRREKVHLEBTIGEIQVTLNKKQKEV---QQL 2176
Qy      992  REQLKVLTAENTMLTSLK 1009
Db      2177  QENLDSTVTQLAAFTKSM 2194

RESULT 12
A61231
myosin heavy chain nonmuscle form A - human
N;Alternate names: cellular myosin heavy chain; myosin type 9; NMHC-A
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Homo sapiens (man)
C;Date: 12-May-1994 #sequence revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: A61231; A34876; I52562; I61692
R;Simons, M.; Wang, M.; McBride, O.W.; Kawamoto, S.; Yanakawa, K.; Gdula, D.; Adelstein,
Circ. Res. 69, 530-539, 1991
A;Title: Human nonmuscle myosin heavy chains are encoded by two genes located on different
A;Reference number: A61231; MUID:91316803; PMID:1860190
A;Accession: A61231
A:Molecule type: mRNA
A;Residues: 1-715 <SIM>
A;Cross-references: UNIPROT:P35579; UNIPROT:Q9UMJ0; GB:M69180; NID:g189029; PIDN:AAA6176
R;Saiz, C.G.; Myers, J.C.; Shows, T.B.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 87, 1164-1168, 1990
A;Title: Human nonmuscle myosin heavy chain mRNA: generation of diversity through altern
A;Reference number: A34876; MUID:90138958; PMID:1967836
A;Accession: A34876
A:Molecule type: mRNA
A;Residues: 715-1961 <SAE>
A;Cross-references: GB:M31013; NID:g189035; PIDN:AAA36349.1; PID:g189036
R;Toothaker, L.E.; Gonzalez, D.A.; Tung, N.; Lemons, R.S.; Le Beau, M.M.; Arnaout, M.A.;
Blood 78, 1826-1833, 1991
A;Title: Cellular myosin heavy chain in human leukocytes: isolation of 5' cDNA clones, c
A;Reference number: I52562; MUID:92003925; PMID:1912569
A;Accession: I52562
A;Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A;Residues: 1-52, 'ENI', 56-659, 'T', 661-868, 'T', 870-930, 'C', 932-1239, 'KG', 1242-1337 <RES>
A;Cross-references: GB:M81105; NID:g188988; PIDN:AAA5988.1; PID:g553596
R;Emmett, W.M.; Haesoon, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.
Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994
A;Title: Identification and overlapping expression of multiple unconventional myosin gen
A;Reference number: A55758; MUID:94294418; PMID:8022818
A;Accession: I61692
A;Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A;Residues: 182-218 <BEM>

```


A;Cross-references: GB:L29141; NID:g457249; PIDN:AAA20904.1; PID:g531134
C;Genetics:
A;Gene: GDB:MYH9
A;Cross-references: GDB:120216; OMIM:160775
A;Map position: 22q12.3-22q13.1
C;Superfamily: myosin heavy chain; myosin motor domain homology
F;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; nucleotide binding site
F;784-764/Domain: myosin motor domain homology <MYO>
F;174-181/Region: nucleotide-binding motif A (P-loop)
F;552-565/Region: actin binding #status predicted
F;626-640/Region: actin binding #status predicted
F;837-1938/Domain: coiled coil #status predicted <COI>
F;837-1277/Domain: S2 #status predicted <DS2>
F;1278-1961/Domain: light meromyosin #status predicted <LM>
F;1939-1961/Domain: carboxyl-terminal <CBT>
F;125/Modified site: N6,N6-trimethyllysine (lys) #status predicted
F;180/Binding site: ATP (lys) #status predicted
F;694,704/Active site: Cys #status predicted

Query Match 6.0%; Score 310; DB 1; Length 1961;
Best Local Similarity 20.6%; Pred. No. 8.3e-06;
Matches 223; Conservative 192; Mismatches 444; Indels 222; Gaps 43;

QY 60 ICGSSIVGMLQNVDFVFAEDHIGITA-----ERYAARGVNYIHQLEHRLPK 113
DB 895 LCAEAELRLATKQQL-BEICHDLARVEEERY-----QHLQAEKKKQ 943
QY 114 NPQNTNPGTSTGTDEAPLAETPTDAESLLEKTPDEAARLVGTSKATQICGKATSG 173
DB 944 NIQLEEQ-----LEBESARQQLQLEKVTTEAKLKLLEEQI-ILEDNCKLAKKLL 999
QY 174 KFEQST-----BETPKILRPKTESKFPKPKERSKITWEKTSVKTCAVGTPTN 228
DB 1000 RIAEFTTNLTBEESKSLAKKNKHEAMITDEERLR-----EKKQROLEK-TRR 1051
QY 229 KTEVLEKGTNNIA-----CPTKETSTKASTNVDSVVEPIFSLFGTRTLEN 275
DB 1052 KLEGSTLSDQIAELQALQALQALQALQALQALQALQALQALQALQALQALQALQAL 1110
QY 276 SQCTKVEEDFNATKIIKSAQNTCLPDATYQKDIKTIHKKIDQWFP8-----ESK 329
DB 1111 SQISELQEDLE-----SERASRNKAEKQKRLDGELEALKTELEDTLDDTAAQOELSK 1164
QY 330 REEDSEYSDSGSLPSSAKTQVCIPESMYQKWEINREVELEPKPSAFKPAVEMQK-T 388
DB 1165 REQEVNI-----LKKTLEEEAKTHEAQIEMRQKHSQAVELEAQLEQVKNLEKAKOT 1221
QY 389 VPNKAFELNEQTLRAAQMFPSKQKDEENSWDSGPC-ETVSQKDVYLPKATH-QKE 446
DB 1222 LENERCELANEVKKVLLQGRDSEHKKKVEAQLEQVKNLEKAKOT 1281
QY 447 FDTLSKLESPVKDGLLKPTCGKRVSLPNKALELKDRETFKAEKSPDKDGLLKPTCGKRV 506
DB 1282 LONVTGLLSQSDSKSKL-----TKDPSALESLQDTQELLQENRKL 1325
QY 507 SLPNKALELKD-----RETLKAEKSPNDGLLK-----PTCGKRVSLPNKALE-----LKDRE 553
DB 1326 SUSTKUKQVEDEKNSFQLEEEBAAKINLEKQATLHAQVADMKKQWEDSVGCLTAE 1385
QY 554 TFKAQMFPS-----SKQKDEENSWDPESFLETLLQN--DVCLPKATH-----QK 598
DB 1386 EVKRLQKDLGLSQRHEKVAAYDKLETKTRLQOELDDLELDHQRQSAACNLEKQK 1445
QY 599 EFD-----TLGKLESPDKGLLKPTCGMKISLPNKALELKORETFKAEKDVSVEST 651
DB 1446 KFDQLLAEEKTISAKYAEERD-----ABEAREKETKALSARALEDA 1489
QY 652 FSLFGPTTNSOSTKVEEDFNLTKEGATKVTGQOERDIIIGIERAPQDQNTKMPSTSEL 711
DB 1490 MEQKAELERLNKQFRTEMEDL-MSSKDDVGVKSV-----HELEKSKALQVVEEMKTLQ 1543
QY 712 GRKEDTKSTSDSEI-ISVSDTONYECLEPATYQKEIKTNGTKIESPEKPSHF-----FP 765

1544 ELEDELQATEDAKRLLEV-----NLQAM-KAOPFERDLQ-----GRDEQSEHKKKQLVQVREM 1595
QY 766 ATQMNSVNVKGLWKNKQTLBADSTTLTKILDA-----LPSCERGR 807
DB 1596 EAELEDERKORSMAVAARKKLEMDLKDLEAHIDSANKRDEAIKQLRKLQAKQKCMREL 1655
QY 808 ELKKNDCOEITA-----KMEQMKKFCVLQKLESEAKEIKSQLENQKAKWEQELCSV- 859
DB 1656 DQTRASREILQAQAKENKELKSMEEAEMQLQEBELAAERAKQAQOERDELADEANSS 1715
QY 860 ---RLPLNQEERKRNVDILKSKI-RPERQLKKLEVKHQLEQTLRIQIDIELKSVTSNIN 915
DB 1716 GKALAL-----LEARIAQLEEELEEQNTELEINDRLKCANLQIDQINADLN 1768
QY 916 -QVSHTHSENDLFHENCMLKKEIAMLKLEVATLKHQOVKKNKYPEDIKILQEKNAELQ 974
DB 1769 LERGHAKQKNEN-----ARQLEKQNKELVK-----LOMEGTVKSKASITALEAKTQLE 1821
QY 975 MTL-----KLKQKTV-----TKRASQYREQLKVLTAENTMLTSKUL 1010
DB 1822 EQLDNETKERQAACKQVRTEKKLKDVLQVDDERRNABQYKQADKASTRLKQLKROLE 1881
QY 1011 E 1011
DB 1882 E 1882

RESULT 13
152300
N;Alternate names: gcp372
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: 152300
R;Sonda, M.; Mismi, Y.; Fujiwara, T.; Nishio, M.; Ikehara, Y.
Biochem. Biophys. Res. Commun. 205, 1399-1408, 1994
A;Title: Molecular cloning and sequence analysis of a human 372-kDa protein localized in
A;Reference number: 152300; MUID:95100974; PMID:7802676
A;Accession: 152300
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-3225 <RES>
A;Cross-references: UNIPROT:Q14789; GB:D25542; NID:g662389; PIDN:BAA05025.1; PID:g808869
C;Superfamily: Giantin

Query Match 6.0%; Score 310; DB 2; Length 3225;
Best Local Similarity 19.3%; Pred. No. 1.5e-05;
Matches 235; Conservative 211; Mismatches 426; Indels 346; Gaps 48;

QY 14 VQNKASUTPILLAIQKSKQKOTVEPLLTKNANANAFNE---SKTALMLAICGSSEIVGM 70
DB 1067 LQDKTNQIDLLQAEISENQAIQKLTSTNDASDGSVALVKETVVISPPCTGSSEHWKP 1126
QY 71 LLOQNV---DVPAEDIHGITAERYAAAGVNYIHQLEHRL-----KLKPNQNTNPGTS 124
DB 1127 ELEBKILALEKEQKQ 1185
QY 125 TGTDEAPLAETPTDAESLLEKTPDEAARLVGTSKATQICGKATSGKPEQTEETPR 184
DB 1186 ---DEQSKENENIGDQLRQLQIQVRESIDGKLPSDQESCS---SSTPGLPEELFKATEQ 1239
QY 185 KILRPTKTSKFSWPAKERSKITWEKTSVKTCEVAGVTPNKTEVLEKGT8-NMIAC 243
DB 1240 HHTQPVLESNLCPDWPMSHSD-----ASALQGGTSVAQIKA 1275
QY 244 PTKE-TSTKASTNVDSVVEPIFSLGTRTENSQCTKVEEDFNATKIIKSAQNTYTC 302
DB 1276 QKLEIAEKVLEKLVST-----TSELKKSSEVQLQEQINKQGLE-----1318
QY 303 LPDATYQKDIKTIHKKIE-----DQWFPESKREE 332
DB 1319 -----ISLTKVSHAEVHAESLQKLESSQLQIAGLEHLRELQPKLDELQKLSKKE 1372


```
Qy 333 DEEYSDGSGSPSSA-----KTQVCIPESMY-----QKVMINEIVEEL 372
Db 1373 DVSYL--SGQLSEKAALTKIETIEIQEDLTKALHTOLEMQAKEHDEHRIKQLQVLCCEM 1430
Qy 373 PEKP-----SAPPAVE--NQKTVPNKAPFKNEQTLR-----AAQMPFSESQKODDE 419
Db 1431 KQKEEIEEERAKQOIQRKLOALISKEALKENKSLQBELSLARGTIERLTLSLADVE 1490
Qy 420 NSWSESPCTVSKQDVVLPK-ATHQKEFDFTLSGKLEBSPVKDGLLKPCTCGRKVSLPNKA 478
Db 1491 SQVSAQN-----KEKDTVLGRALLQBERDKLITEMDRSLLENQSLSSSC-----ES 1537
Qy 479 LELKDRFTFKAESPDGGLKPTCGRKVSLPNKALELKDOR----- 518
Db 1538 LKL-----ALEGLTDEKELVKEIEISLSSKSKIAESTEMOEKHELQKEYEILLQSVENVSN 1593
Qy 519 -----ETLKAESPNDGGLKPTCGRK-----VSLPN 544
Db 1594 EAERIQHVVAVROEKQELYKLRSTEANKKETEKQLQAEQBEEMEKWKRFKSKQQ 1653
Qy 545 KALELKORETFKAAQMPFSESQKQDDSENSWDFESFLETLQNDVCLPKATH--QKEPDT 602
Db 1654 KILELEENDRLRAEVHPAGDTAKE-----CMETLSSNASMKEELERVKMEVET 1703
Qy 603 LSGKLEE--SPDKDGLLKPCTCGMKISLPN-----KALELKDRFTFKAE--VS 646
Db 1704 LSKKFQSLMSKSLSEEVQDLKHQIEGNVSKQANLEATEKHNDQNTVTEBGTQSIPGET 1763
Qy 647 SVESFSLFGKPT--TENSQSTK-----VEEDFNLTTK--EGATKVTGQOERDGIIBRAP 699
Db 1764 BEQDSLGMSTRPTCSSEVPSAKSANPAVSKDFSSHDEINNYLQIQIDQKERIAGLEE--- 1820
Qy 700 QDQTNKMPSTSELGRKEDTK-----STSDSEIISVSDTQNYECLPEATYQKEIKTTNGKIEE 755
Db 1821 EKQKNKBPSTLENEKNTLLSQISTQKQELQWLOEVTYKNNLLNQIQEELSRVT--KLKE 1879
Qy 756 SPEKSPHEPATENQNSVPNKGLEBWKNNQTLRADSTTSLKILDALPSCERGLKONCE 815
Db 1880 TAE-----EEKDLEERLMNQ-----LAELNGSI-----GNVCQ 1908
Qy 816 QIT---AKMEQMKNFVLOKELSENAEKISOLENQAKWQELCSVRL-----PL 863
Db 1909 DVTDAQIKNELESEMKNLKKVCSELEBEKQOLVKEKTKVSEIRKEYLEKIQGAQKEPG 1968
Qy 864 NOBEKERNVDILKEIRPEBQRLKLEVKHQ-----LEQTLRIODI-----ELKSVT 911
Db 1969 NKSHAKELQ-ELLKEKQEVQLOKDC--IRYQEKISALERTVKALBFPVOTESQKDLKITK 2026
Qy 912 SNLQO-VSHTHESNDL-----FHENCMLKXGI-----AMLKLEV 945
Db 2027 ENLAQAVEHRKKAQAEALASFVLLDDDTQSEAAARVLADNLKLKXELQSNKESVKSMQKQD 2086
Qy 946 ATLKHQHVKNKYPEDIKILQEK-----NAELMTLKLKQKTVTKRASQY 991
Db 2087 EDLERLEQAEERKLEKKNQWQELDALRREKVVHLEETIGTIGVTLNKKQKEV-----QOL 2142
Qy 992 REQQLVLTAEMLTSLK 1009
Db 2143 QENLDSITVQLAAFTKSM 2160

RESULT 14
T14156
kinasein-related protein - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T14156
R/Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A/Title: CNRP-E is a plus end-directed kinetochore motor required for metaphase chromoso
A/Reference number: Z17893; MUID:98028574; PMID:9363944
A/Accession: T14156
A/Status: preliminary; translated from GB/EMBL/DBJ
```

A:Molecule type: mRNA

A:Residues: 1-2954 <WOO>

A:Cross-references: UNIPROT:O42263; EMBL:AF027728; NID:g2586070; PID:g2586071; PIDN:AAAC

C:Genetics:

A:Gene: XSNP-E

C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match 5.9%; Score 307.5; DB 2; Length 2954;

Best Local Similarity 19.9%; Pred. No. 1.8e-05;

Matches 245; Conservative 200; Mismatches 402; Indels 383; Gaps 53;

```
Qy 17 KASLTPLLLAIQKSKQTV-----FLTYKANANAFNESCKTALMLAICE 62
Db 1211 KGHLTDSQLSIEKLEQLENLEVTETLQTLQEBMKNTIERNELQTNFEDLK----- 1260
Qy 63 GSSIVGMLLQONVDVFAEDIHGITAERYAAARGVNYHQOLLEHIRK-----LP 112
Db 1361 AEHDSLQDLSENEIQSIE-----TQDELRAAQBELREQOLQVDSFRQOLLDSCVGISSP 1315
Qy 113 KNQNTNPEGTSTGTPDEAAPLAERTPTDAESLLEKTPDEAARLVEGTSAKIQCLGKATS 172
Db 1316 NHDAVANQEKVSLG-----EVNSL-----QSEMLRGERDELQTSKALV 1354
Qy 173 GKPEQSTETPRKILRPTKETSEKSPAKERSRKITWEEKETSVKTECVAGVTPNKTEV 232
Db 1355 SELE-----LLRAHVKSVEGEN--LEITTKLANGLEKEILGKSE-----ESEV 1394
Qy 233 LEKGTSNMIACTKETSTKASTNV--DVSSVEPIFSLFQTRTIENSQCTKVEEDFNLATKI 291
Db 1395 LKSMLENL-----KEDNNKLKQAEYSSENQPSL-----BEVFGSQKL 1435
Qy 292 ISK-----SAAQNTYTCLPDATY-----Q 309
Db 1436 VDEIEVLKAQKAAEERLEIKDRDYFELVQNTANLVGKLETPLQADHEEDSIDRRSEE 1495
Qy 310 KDIKTIHKKIEDQMPFSESKREDEEYSDGSGSPFESSAKTQVCIPESMYQKVMINEIREV 369
Db 1496 MEIKVLGKLERNYLLERLQEEKLELS--NKLLEIKEMETSVLLKDDLOQKJESLSEN 1554
Qy 370 BELPEK--PSAPK---PAVEMQKTVPNKAPFKNEQTLRAAQMPFSESQKQDDSENSWDS 424
Db 1555 IILKENIDITLKHSHDTPQAQLOKT--QOELQAKNLALAAASDNCPI-----TQEKETSADC 1608
Qy 425 ESPCETVSQKDVVLPKATHQK--EFDTL---SGKLEESPVK-----DGLLKPTCGRKVSL 474
Db 1609 VHPLE---EKILLTTEELHQTNEQKELHKEKLEQAQVELKCEVEHLMKSMIESKSSL 1665
Qy 475 PNKALELKDRB-----TFKAESP 492
Db 1666 ESLQHEKHDTQOALLALKQOMQVVTQBKKELOQTHEHLTAEDVHLKENIELGLNFKNEAQ 1725
Qy 493 DKDGLLKPTCGRKVSLPNKALE-----LKDR-----ETLKAESPNDGGL 532
Db 1726 QK--TTKEQC---LINENKELEQSHRLQCBIEBLMSKLDKDKESALETLKES----- 1773
Qy 533 KPTCGRKVSLPNKALE--LKDRFTFKAAQ--MFPSESKQKDDSENSWDFESFLET----- 583
Db 1774 -----QKVINLQEMWVMEBELKNSQRTVIAERDQLOQDRLRES--VENSLETQDDLR 1826
Qy 584 -----LQONDVCLPKATHQKEPDTLSGKLESPDKG 615
Db 1827 KAQBALQQQDKVQBELTSQISVLQEKISLLENQMLYNAVTVK-----ETLSERDDLNQSKQH 1883
Qy 616 LLKPTCGMKISLPNKALELKDRFTFK-----EDVSSVSESTF-----SLFG 656
Db 1884 LFSBIETLSLSLKEKBPALQAEKDKDAARKTTIDITEKISNIEBQLLQQAUNUKETIYE 1943
Qy 657 KPT-----TENSQSTKVEEDFNLTTPKEGATKVTVTCQOERD-----IGIER-- 697
Db 1944 RESLIQCKEQALANTEHLRETLKSKDLALGKWE-----QERDEAANKVIALTERMS 1994
Qy 698 APQDQTNKMPST-----ELGRKE-----DTKSTSEIISVSDTQNYECLPEATYQKEIKT 748
```

Db 1995 SLEEQINENVTLKEGGEKETFYLRQPSQOQSSQMEELRESLTKYDLOLEAEKEISE 2054

Qy 749 TNGIIESPEKSPHFEPATEMNSVPNKGLEWKNKOTLRADSTYLSKILDALPSCERGR- 807

Db 2055 ATNEIKNLTKISLEEEILQNASILNEAV--SERENLRHSKQOLVSELEQLSLTKSRD 2112

Qy 808 ----ELKDNCEQITAKHQWKNKFCVLQKELSEAKIQLQENOKAKWQELCSVRLPL 863

Db 2113 HAPAQSKREKDEAVN-KIASLAEBIKILTKEMDEFKDSKESLQSQSHLSEELCTYKTEL 2171

Qy 864 NQEEKERNV-DILKEKIRPEQQRKKL-EVKHQLEQTLRIQDTELKSVTSNLMNOVSHTH 921

Db 2172 QMLQKQEKEDINNKLAKEKVEDELLQHLSSLEQLDQI-----QNELR-----NEKLRYN 2221

Qy 922 ESENDLPHENCKMEKELAMLKLEVATLKHOHQVKNKYFEDIKILQKKNBLQMTLKLKQ 981

Db 2222 E----LCEKMDIMEKEISVLR-----MQNEPQOEEDDVAERMIDLESRNQEI-----Q 2266

Qy 982 KTVTKRASQYREQLKVLTAENTMLTSKLKE 1011

Db 2267 ELMEXISAVYEQHTLJSS-----JSSLEQK 2292

RESULT 15

JC5837

364K Golgi complex-associated protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 05-Mar-1998 #sequence_revision 13-Mar-1998 #text_change 09-Jul-2004

C:Accession: JC5837

R:Toki, C.; Fujiwara, T.; Sohda, M.; Hong, H.S.; Misumi, Y.; Ikehara, Y.

Cell Struct. Funct. 22, 565-577, 1997

A:Title: Identification and characterization of rat 364-kDa Golgi-associated protein rec

A:Reference number: JC5837; PMID:98093490; PMID:9431462

A:Accession: JC5837

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3187 <TOK>

A:Cross-references: UNIPROT:Q63714; DDBJ:D25543; NID:9516825; PIDN:BAA05026.1; PID:95168

C:Comment: This protein plays a role in the formation and maintenance of the characteris

C:Superfamily: giantin

F:49-549,624-1176,1238-1707,1763-3114/Domain: coiled-coil leucine zipper #status predic

F:3165-3187/Domain: membrane anchor #status predicted <MAD>

Query Match 5.8%; Score 302.5; DB 2; Length 3187;

Best Local Similarity 21.1%; Pred No. 3.2e-05;

Matches 253; Conservative 197; Mismatches 451; Indels 297; Gaps 54;

Qy 1 MVATLLSYGAVIEVQNK-----ASLTPLLLAIQKRSKOTVEFLTKRANANAFNESHCTALM 57

Db 1453 LTKSLADVESQVSQVQKQKXKALLGLKALLQERDKLIVE--MDKSLLENQSLGSCESLK 1510

Qy 58 LAICGSSSIVGMILQONVDVFAEDIHGITAERYAAAGVNVYHQOLLEHIRKLPKNQON 117

Db 1511 LAL--GG-----LTDEKELMKELKESVRCISKIESTQWQKHKELOKEYEVLLQSTYEN 1561

Qy 118 TNPGEFSTGTPDEAALERTPTDABSLLEKTPDPAARLVGTSKIAQICLGKATSKFQ 177

Db 1562 VSNE-----ABRIQHVSVQKQEVYAKLSAESDKREKQLQDA--EQ 1606

Qy 178 STEETPRKILRPTKTSKPF-----SWPAYERSK 207

Db 1607 EMEEMKVKRKFQKQKILELEENRDLRAEQPVGGANESMEALLSSNASLKEELER 1666

Qy 208 IITWEEKETSVKTECVAGVTPNKTVELEKGTNN-----IACPTKETSTKASTNVDS 259

Db 1667 ITLEYTKLSEFFALMA-----EKNTLSEETRNKLQVQAEQLQASLETETKSDPKDV- 1721

Qy 260 SVPEPISLFCFTRTIEN---SQCTKVEE-DFNLATKIISKSAQNYTCLPDA-TYQKDIKT 314

Db 1722 -IEEVTEAVVGKQEQDSLSSENAKLEDAEATLLANSAPGVSETFFSHDDINNYLQQLDQ 1780

Qy 315 INHKIEDQMPPSKESKEDEBEYSW-----DSGSLFPSSAK-TQVCIPESMYQKWEINRE 368

Db 1781 LKGRIAE-----LEMEKQKQDRELSQTLNEKFNALLTQISAKOSELKLEBEVAKINMLNQO 1836

Qy 369 VBELEPKSPAPPAVE-----NQKTPVNKAPEL-----KNEQ----- 400

Db 1837 IQEELSRVTKLETAEEDDEEDLEERLMNQALNELSGISNYQYQDVTDAQIKNEQLSEMQN 1896

Qy 401 TLRAAQMFPSQKQDDEENSDSPCTVTSQKQVYLPK-----ATHQKEFDTL 450

Db 1897 LKRCVSELEBEKQOLVKEIKTVESEIRKE-----YMEKIQAQKGPQSHAKELQEL 1949

Qy 451 SGKLESPPVKDGLLKPTC-----GRKVSIPN--KALELKDRETFFKAESPDKDGLLKPTCCR 504

Db 1950 ---LKEQEQEVQLOQDCIRYLGRISALEKTVKALFEVHTES-----QKD--LDAIKG- 1997

Qy 505 KVSILPNKALELKDRETLKAEPSDNDGLLKPTCGRKVSLPNKALELKDRETFKAAQFPSE 564

Db 1998 ----NLAQAVEHHKQAQELSFKILLDDTQSEAAARVLADNLKLK- KELSQNKESIKSQ 2051

Qy 565 SKQKDEENSDWDFLETLQNDVCLPKATHQKEFDTLSGKLESPPDKDGLLKPTCG-M 623

Db 2052 IKQKD-----EDLLRLLEQ-----ABEKIRKEKKNMQEKLDAHLREKAHVEDTLAEI 2098

Qy 624 KISILPNKALELKDRETFKAEADVSSVESTSLFGKPTTENS-----QSTKVEEDF 672

Db 2099 QVSLTRKDKDMKELQ-----QSLDSTLAQLAAFTKSMSSLODDRDRVIDEAKKWEORF 2151

Qy 673 NLTTKEGATVTVGQOERDIGIYER---APQDOTNRNPTSELGRKEDTKTSDSEIISVS 729

Db 2152 G-----DAIQT-----KEEVRLAKEENCTALKDQLRQWTI-----HMBELK-----ITVS 2191

Qy 730 ---DTQNYECLPEATYQKEIKITTKNGKIBESPEKPSHPPEPATMNSVPNKGLEWKNK-Q 784

Db 2192 RLEHDKIEWESKQATELHQHQKAYDKLOEENKELMSQLEAGQLYHDSKRNELTKLESELK 2251

Qy 785 TLRADSTTLKILDALPSCERGRE--LK-----KDNCEQ-----ITAKME----- 822

Db 2252 SLKQDQSTDLKNSLEKCKREHNNLEGIQKQEAQIONCKFCEQLETLDTLTASRELTTLRHD 2311

Qy 823 -----QMKNKFCVLQKELSE-----AKEIKSQLENQKAKWQELCSVRLPLNQE 866

Db 2312 EINVKEQKIISLLSGKEEAQVAIAELHQHSKEIK-ELENLLSQEBEE-----NLTL 2364

Qy 867 BEKRNVDILKEKIRPEQQRKK-LEVKHQLEQTLR-----IQDIELKSVTSNLMNOVSHTH 921

Db 2365 EENKRAVEKTNQLTEALETIKKESLEQKQALDSFVKSMSSLOD--DRDRIVSDYRQLSERH 2423

Qy 922 -----ESENDLPHENCKMEKELAML-----KLEVATLKHOHQVKNKYFE 961

Db 2424 LSVILEKDELIQDAAAENNKLEIRGLRGHMDLNSENAKLDAELIQYRDLNEVITIK 2483

Qy 962 DIKILQENKNAELQMTLKLKOKTV-----TKRASQYREQ-----LKVLTAEINTMTLSKLK 1010

Db 2484 DSQQRQLLEAQLQONKELRNECVKLEGRKLGSEAKQSLQMSLDALQFENQGLSKK 2541

Search completed: February 19, 2005, 02:12:27

Job time : 38.5932 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	5173	100.0	1011	16	US-10-408-765A-1557	Sequence 1557, App	
2	5173	100.0	1239	13	US-10-007-805-577	Sequence 577, App	
3	5173	100.0	1239	14	US-10-076-622-577	Sequence 577, App	
4	5173	100.0	1239	14	US-10-124-805-577	Sequence 577, App	
5	5156	99.7	1225	14	US-10-177-293-334	Sequence 332, App	
6	3075	59.4	1341	13	US-10-007-805-565	Sequence 565, App	
7	3075	59.4	1341	14	US-10-076-622-565	Sequence 565, App	
8	3075	59.4	1341	14	US-10-177-293-334	Sequence 334, App	
9	3075	59.4	1341	14	US-10-124-805-565	Sequence 565, App	
10	3075	59.4	1341	15	US-10-058-270A-4	Sequence 4, Appli	
11	3069	59.3	1349	13	US-10-007-805-573	Sequence 573, App	
12	3069	59.3	1349	14	US-10-076-622-573	Sequence 573, App	
13	3069	59.3	1349	14	US-10-124-805-573	Sequence 573, App	

121 QY EGTSTGTPDDEAALPRTPTDPAESLLEKTPDEAARLVESGSAKIQCLGKATSGKFEOSTE 180
121 DB EGTSTGTPDDEAALPRTPTDPAESLLEKTPDEAARLVESGSAKIQCLGKATSGKFEOSTE 180
181 QY ETPRKILRPRTKTESKFSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 240
181 DB ETPRKILRPRTKTESKFSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 240
241 QY IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAKIIISKAANY 300
241 DB IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAKIIISKAANY 300
301 QY TCLPDATYQDKITINHIEQDMPFSESKREDEEYSWDSGLPESAKTOVCIPESMYQ 360
301 DB TCLPDATYQDKITINHIEQDMPFSESKREDEEYSWDSGLPESAKTOVCIPESMYQ 360
361 QY KWEINREVEELPEKPSAFKPAVEMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDEEN 420
361 DB KWEINREVEELPEKPSAFKPAVEMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDEEN 420
421 QY SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
421 DB SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
481 QY LKDRETFKAESPDKDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPTCGRKV 540
481 DB LKDRETFKAESPDKDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPTCGRKV 540
541 QY SLPNKALELKDRETLKABSPNDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPT 600
541 DB SLPNKALELKDRETLKABSPNDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPT 600
601 QY DTLGKLEESPDGKGLLKPTCGMKISLPNKALELKDRETLKABSPNDGLLKPTCGRKV 660
601 DB DTLGKLEESPDGKGLLKPTCGMKISLPNKALELKDRETLKABSPNDGLLKPTCGRKV 660
661 QY ENSQSTKVEEDFNLTKEGATKTVTGOQERDGIIERAPOQOTNKMPSELGRKEDTKST 720
661 DB ENSQSTKVEEDFNLTKEGATKTVTGOQERDGIIERAPOQOTNKMPSELGRKEDTKST 720
721 QY SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSHFEPATEMNSVPNKGLEW 780
721 DB SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSHFEPATEMNSVPNKGLEW 780
841 QY IKSQLENQAKWEQELCSVRLPLNQEEKRRNVVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
841 DB IKSQLENQAKWEQELCSVRLPLNQEEKRRNVVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
901 QY RIQDIELKSVTNLQVSHSENDLPHENCMLKKEIAMLKLEVATLKHQHVQENKVF 960
901 DB RIQDIELKSVTNLQVSHSENDLPHENCMLKKEIAMLKLEVATLKHQHVQENKVF 960
961 QY EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLVLTAEINTMLTSKLE 1011
961 DB EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLVLTAEINTMLTSKLE 1011

RESULT 2
US-10-007-805-577
; Sequence 577, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hegpler, William T.

; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 577
; LENGTH: 1239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-577

Query Match 100.0%; Score 5173; DB 13; Length 1239;
Best Local Similarity 100.0%; Pred. No. 8.4e-263;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVATLSYGAVISVQNKASLTPLLLAIQKESKOTVEFLTKNANANAFNESCCTALMLAI 60
DB 1 MVATLSYGAVISVQNKASLTPLLLAIQKESKOTVEFLTKNANANAFNESCCTALMLAI 60
QY 61 CEGSSEIVGMLLQONVDVFAEDIHGITAERYAAARGVNYIHQOLLEHRLKPKNPQNTNP 120
DB 61 CEGSSEIVGMLLQONVDVFAEDIHGITAERYAAARGVNYIHQOLLEHRLKPKNPQNTNP 120
QY 121 EGTSTGTPDDEAALPRTPTDPAESLLEKTPDEAARLVESGSAKIQCLGKATSGKFEOSTE 180
DB 121 EGTSTGTPDDEAALPRTPTDPAESLLEKTPDEAARLVESGSAKIQCLGKATSGKFEOSTE 180
QY 181 ETPRKILRPRTKTESKFSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 240
DB 181 ETPRKILRPRTKTESKFSWPAKERSRKITWEEKETSVKTECVAGVTNPKTEVLEKGTSM 240
QY 241 IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAKIIISKAANY 300
DB 241 IACPTKETSTKASTNDVSSVEPIFSLFGTRTIENSQCTKVEEDFNIAKIIISKAANY 300
QY 301 TCLPDATYQDKITINHIEQDMPFSESKREDEEYSWDSGLPESAKTOVCIPESMYQ 360
DB 301 TCLPDATYQDKITINHIEQDMPFSESKREDEEYSWDSGLPESAKTOVCIPESMYQ 360
QY 361 KWEINREVEELPEKPSAFKPAVEMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDEEN 420
DB 361 KWEINREVEELPEKPSAFKPAVEMQKTVPNKAFELKNEQTLRAAQMPFSESKQKDEEN 420
QY 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
DB 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
QY 481 LKDRETFKAESPDKDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPTCGRKV 540
DB 481 LKDRETFKAESPDKDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPTCGRKV 540
QY 541 SLPNKALELKDRETLKABSPNDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPT 600
DB 541 SLPNKALELKDRETLKABSPNDGLLKPTCGRKVSLPNKALELKDRETLKABSPNDGLLKPT 600
QY 601 DTLGKLEESPDGKGLLKPTCGMKISLPNKALELKDRETLKABSPNDGLLKPTCGRKV 660
DB 601 DTLGKLEESPDGKGLLKPTCGMKISLPNKALELKDRETLKABSPNDGLLKPTCGRKV 660
QY 661 ENSQSTKVEEDFNLTKEGATKTVTGOQERDGIIERAPOQOTNKMPSELGRKEDTKST 720
DB 661 ENSQSTKVEEDFNLTKEGATKTVTGOQERDGIIERAPOQOTNKMPSELGRKEDTKST 720
QY 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSHFEPATEMNSVPNKGLEW 780
DB 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSHFEPATEMNSVPNKGLEW 780

Db 121 EGTSTGTPDEAALPAAERTPDTAELSLEKTPDEAARLVGTSKTIQCLGKATSGKFEQSTE 180
Qy 181 ETPRKILRPKTSEKTSWPAKERSRKITWEEKTSVKTECVAGVTNPKTEVLEKGTSM 240
Db 181 ETPRKILRPKTSEKTSWPAKERSRKITWEEKTSVKTECVAGVTNPKTEVLEKGTSM 240
Qy 241 IACPTKETSTKASTNDVSSVEPIFSLFGTTRIENSQCTKVEEDFNATKIISKSAQNY 300
Db 241 IACPTKETSTKASTNDVSSVEPIFSLFGTTRIENSQCTKVEEDFNATKIISKSAQNY 300
Qy 301 TCLPDATYQKDIKTINHKEIDQMPSESREDEEYSWDSGLSFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQKDIKTINHKEIDQMPSESREDEEYSWDSGLSFESSAKTQVCIPESMYQ 360
Qy 361 KWEINREVEELPEKPSAFKPAVENQKTPVNPNAFELKNEQTLRAAQMPFSSKQKDDEN 420
Db 361 KWEINREVEELPEKPSAFKPAVENQKTPVNPNAFELKNEQTLRAAQMPFSSKQKDDEN 420
Qy 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSPLNKALE 480
Db 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSPLNKALE 480
Qy 481 LKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGKRV 540
Db 481 LKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGKRV 540
Qy 541 SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALE 600
Db 541 SLPNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALE 600
Qy 601 DTLSGKLEESPVKDGLLKPTCGRKVSPLNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALE 660
Db 601 DTLSGKLEESPVKDGLLKPTCGRKVSPLNKALELKDRETFKAESPDKDGLLKPTCGRKVSPLNKALE 660
Qy 661 ENSQSTKVEEDFNATKIISKSAQNY 720
Db 661 ENSQSTKVEEDFNATKIISKSAQNY 720
Qy 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSFEPATQNMPTSELGRKEDTKST 780
Db 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIESPEKPSFEPATQNMPTSELGRKEDTKST 780
Qy 781 KNOQTIRASTTILSKILDALPSCERGERELKDNCEQITAKMEQKMKFVLOKELSEAKE 840
Db 781 KNOQTIRASTTILSKILDALPSCERGERELKDNCEQITAKMEQKMKFVLOKELSEAKE 840
Qy 841 IKSQLENOQAKWQELCSVRLPLNQEEKRRNVNDELKEKIRPEEQRLKLEVKHOLEQTL 900
Db 841 IKSQLENOQAKWQELCSVRLPLNQEEKRRNVNDELKEKIRPEEQRLKLEVKHOLEQTL 900
Qy 901 RIQDIELKSVTSLNQVSHTHSENDELPHENCMKKEIAMLEKLEVATLKHQVKNKYF 960
Db 901 RIQDIELKSVTSLNQVSHTHSENDELPHENCMKKEIAMLEKLEVATLKHQVKNKYF 960
Qy 961 EDIKILOEKNAELQMTLKLKQKVTTKRASQYREQLKVLTAENTMLTSKLKE 1011
Db 961 EDIKILOEKNAELQMTLKLKQKVTTKRASQYREQLKVLTAENTMLTSKLKE 1011

RESULT 5
US-10-177-293-332
; Sequence 332, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lilie, James
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen

; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 332
; LENGTH: 1225
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-332

Query Match 99.7%; Score 5156; DB 14; Length 1225;
Best Local Similarity 99.7%; Pred. No. 6.5e-262;
Matches 1008; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MVATLLSYGAVIEQVONKASLTPLLLAIQKRSKQTVFLLTKNANANAFNESCCTALMLAI 60
Db 154 MVATLLSYGAVIEQVONKASLTPLLLAIQKRSKQTVFLLTKNANANAFNESCCTALMLAI 213
Qy 61 CEGSEISVGLMLQNVDPFAEDIHGITAERYAAAGVNYHQOLLEHRLKPNPQNTNP 120
Db 214 CEGSEISVGLMLQNVDPFAEDIHGITAERYAAAGVNYHQOLLEHRLKPNPQNTNP 273
Qy 121 EGTSTGTPDEAALPAAERTPDTAELSLEKTPDEAARLVGTSKTIQCLGKATSGKFEQSTE 180
Db 274 EGTSTGTPDEAALPAAERTPDTAELSLEKTPDEAARLVGTSKTIQCLGKATSGKFEQSTE 333
Qy 181 ETPRKILRPKTSEKTSWPAKERSRKITWEEKTSVKTECVAGVTNPKTEVLEKGTSM 240
Db 334 ETPRKILRPKTSEKTSWPAKERSRKITWEEKTSVKTECVAGVTNPKTEVLEKGTSM 393
Qy 241 IACPTKETSTKASTNDVSSVEPIFSLFGTTRIENSQCTKVEEDFNATKIISKSAQNY 300
Db 394 IACPTKETSTKASTNDVSSVEPIFSLFGTTRIENSQCTKVEEDFNATKIISKSAQNY 453
Qy 301 TCLPDATYQKDIKTINHKEIDQMPSESREDEEYSWDSGLSFESSAKTQVCIPESMYQ 360
Db 454 TCLPDATYQKDIKTINHKEIDQMPSESREDEEYSWDSGLSFESSAKTQVCIPESMYQ 513
Qy 361 KWEINREVEELPEKPSAFKPAVENQKTPVNPNAFELKNEQTLRAAQMPFSSKQKDDEN 420
Db 514 KWEINREVEELPEKPSAFKPAVENQKTPVNPNAFELKNEQTLRAAQMPFSSKQKDDEN 573
Qy 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSPLNKALE 480
Db 574 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSPLNKALE 633
Qy 481 LKDRETFKAESPDKDGLLKPTCGRKVSPLNKALELKDRETLKABSPDNDGLLKPTCGKRV 540

Db	1112	LKVLAENTMLTSKJKE	1128	
RESULT 7				
US-10-076-622-565				
; Sequence 565, Application US/10076622				
; Publication No. US20030023036A1				
; GENERAL INFORMATION:				
; APPLICANT: Houghton, Raymond L.				
; APPLICANT: Sleath, Paul R.				
; APPLICANT: Persing, David H.				
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY				
; FILE REFERENCE: 210121.470C11				
; CURRENT APPLICATION NUMBER: US/10/076,622				
; CURRENT FILING DATE: 2002-02-13				
; NUMBER OF SEQ ID NOS: 627				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 565				
; TYPE: PRT				
; LENGTH: 1341				
; ORGANISM: Homo sapiens				
US-10-076-622-565				
Query Match 59.4%; Score 3075; DB 14; Length 1341;				
Best Local Similarity 57.4%; Pred. No. 8.6e-153;				
Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;				
QY	1	MVATLLSYGAVTEVQNKASLTPLLLAIQKRSQTVFELLTKNANANAFNESKCTALMLAI	60	
Db	98	VVAKLLSHGAVLEVNKASLTPLLLSITKRSQIVFELLKNANANAVNYKCTALMLAV	157	
QY	61	CGSGSEIVGMLLQONVDVPAEDIHGTAIRYAAAGVNYHIOOLLEHRIKLPKPNQNTNP	120	
Db	158	CHGSSEIVGMLLQONVDVPAEDICGVTAHYAVTCGFHHIHEQIMEYIRKLSKHQNTNP	217	
QY	121	EGTSTGTPEEAAPLA-----ERTPDTSLSLEKTPDEAARLVE	158	
Db	218	EGTSAGTPEEAAPLAERTPDTSLSLEKTPDEAARLVEERTPDTSLSLEKTPDEAARLVE	277	
QY	159	GTSAKIQCLGKATSGKFEQSTETPRKILRPKTSEKFSWP-----	200	
Db	278	GTSKIQCLGKATSGKFEQSAETPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPR	337	
QY	201	-----AKRSRKITWEKETSVKTECVAGVTPNKTVELEKGTSMNIACP	244	
Db	338	EIMSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTGCVARVTSNKTVKLEGRSKOIACP	397	
QY	245	TKETSKASTNVDVSSVEPIFSLFGTRTIENSQCTKVERDFNLATKIISKSAQAQNYTCLP	304	
Db	398	TKESSTKASAN-----	408	
QY	305	DATYQKDIKTIHKBEDQMPFSSKREEDBYSWGSLFESSAKTQVCIPESMTQKWE	364	
Db	409	-----DQRPSPSKQEDDEYSCDSRSLFESSAKIQVCIPESIIYQKWE	452	
QY	365	INREVELPKPSAFPAVEMQTVNKAFELKNEOTLEAAQMPFSESQKQDEENWDS	424	
Db	453	INREVEPPKPSAFPAEMQNSVFNKAFELKNEOTLRADPMFPFSESQKQDEENWDS	512	
QY	425	ESPCETVSQDVLPKATHQKEPDTLSGLKEESPVRDGLLKPTCGRKVSLPNKALELKDOR	484	
Db	513	ESICETVSQDVLPKATHQKEIDKINGKLEESPNDGLLKATCGMKVSIPTKALELKDOR	572	
QY	485	ETPKAESPDKGLLKPTCGRKVSLPNKALELKDRETLKAESPDNDGLLKPTCGRKVSLPN	544	
Db	573	QTFKAEPGKPSAFEP-----ATEMQK-----SVFN	598	
QY	545	KALELKDRETFKAQMPFSSKQKDEENSWDPESFLETLQNDVCLPKATHQKEPDTLS	604	
Db	599	KALELKNEQWRADEILPSESQKQDEENSWDTESETVTSQKDVCLPKAAHQKEIDKIN	658	
QY	605	GKLEESPDKGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESTFSLFGKPTTENSQ	664	
Db	659	GKLEGSFVRDGLLKANKCMKVSIPTKALELMDMQTFKAEPPEK-PSAF-----EPAIEMQK	713	
QY	665	STKVEEDFNLTTEGATKTVTCQOERDGIIEIRAQODQTNKMPTELSGKEDTKSTDSSE	724	
Db	714	SV-----PNKALELKNBQTLRADRILP-----SESKQKQIYESWDSSE	751	
QY	725	IISVSDTONVECLPEATYQKEIKTTNGKIEES-----	756	
Db	752	SICETVSQDVCLPKATHQKEIDKINGKLEESPNDNDGFLKAPCRMKVSIPTKALELMDMQ	811	
QY	757	-----PEKPSHFEPATEMNSVFNKGLKWKKNQTLRA-----	788	
Db	812	TFKAEPPEKPSAFPAIEMQKSVFNKALELKNBQTLRADQMPFSESQKQKVEENSWDSSES	871	
QY	789	-----DSTTSLKILDALPSCERGRRLKKNQCQITA	819	
Db	872	LRETVSQDVCPKATHQKEMDKISKLEEDSTLSKILDTVHSCEARLQKDHCEQRTG	931	
QY	820	KMEQMKKFCVLQKELSEAKEIKSQLENQKAKWEQELCSVRLPLNOEERKRNVDILKKE	879	
Db	932	KMEQMKKFCVLKELSEAKEIKSQLENQKAKWEQELCSVRLTLNOEERKRNADILNEK	991	
QY	880	IRP-----EQLRKLEVKHOLEQTLRIQDIELKSVTSLNQLVSHTHESENDLFHENCML	934	
Db	992	IREELGRIEQRHKELEVKQLEQALRIQDIELKSVESNLNQVSHTHESENDLFHENCML	1051	
QY	935	KKSIAMLEKLVATLKHOHGVKENKYFEDIKLOEKNAELQMTLKKOKTVTKRASQYREQ	994	
Db	1052	KKSIAMLEKLVATLKHOHGVKENKYFEDIKLOEKNAELQMTLKKOKTVTKRASQYREQ	1111	
QY	995	LKVLTAENTMLTSKJKE	1011	
Db	1112	LKVLAENTMLTSKJKE	1128	
RESULT 8				
US-10-177-293-334				
; Sequence 334, Application US/10177293				
; Publication No. US20030124128A1				
; GENERAL INFORMATION:				
; APPLICANT: Lillie, James				
; APPLICANT: Glatt, Karen				
; APPLICANT: Zhao, Xumei				
; APPLICANT: Gannavarpu, Manjula				
; APPLICANT: Kamatkar, Shubhangi				
; APPLICANT: Mertens, Maureen				
; APPLICANT: Myer, Vic				
; APPLICANT: Wang, Youzhen				
; APPLICANT: Xu, Yongyao				
; APPLICANT: Hoersch, Sebastian				
; APPLICANT: Monahan, John				
; APPLICANT: Meyers, Rachel E.				
; APPLICANT: Bast Jr., Robert C.				
; APPLICANT: Hortobagyi, Gabriel N.				
; APPLICANT: Pusztai, Lajos				
; APPLICANT: Sahin, Aysegul				
; APPLICANT: Mills, Gordon B.				
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,				
; FILE REFERENCE: MRI-038				
; CURRENT APPLICATION NUMBER: US/10/177,293				
; CURRENT FILING DATE: 2002-06-21				
; PRIOR APPLICATION NUMBER: US 60/299,887				
; PRIOR FILING DATE: 2001-06-21				
; PRIOR APPLICATION NUMBER: US 60/301,572				
; PRIOR FILING DATE: 2001-06-27				
; PRIOR APPLICATION NUMBER: US 60/306,501				
; PRIOR FILING DATE: 2001-07-18				
; PRIOR APPLICATION NUMBER: US 60/325,002				
; PRIOR FILING DATE: 2001-09-25				
; PRIOR APPLICATION NUMBER: US 60/362,585				


```
Qy 725 IISVSDTONYECLEPEATYQKEIKTTNGKIBES----- 756
Db 752 SLCEVTSQKDVCLPKATHQKEIDKINGKLEESPNDGFLKAPCRMKVSIPTKALELMDMQ 811
Qy 757 -----PEKSHFEPATEMOMQSVNKGLEWKNKOTLRA----- 788
Db 812 TFKAEPPEKPSAFPAIEMQSVNKALELQNEQTLRADQMFPSKQKKVEENSWDES 871
Qy 789 -----DSTTLKILDALPSCGERELKKNCEQITA 819
Db 872 LRETVSQKDVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTG 931
Qy 820 KMEOMKMKFCVLOKELSEAKIKSOLENOKAKWEQELCSVLPNOBEERKNVDILKEK 879
Db 932 KMEOMKMKFCVLLKKLSEAKIKSOLENOKKVKWEQELCSVRLTLNQBEEKERNADILNEK 991
Qy 880 IRP-----EEOLRKKLEVKHOLEQTLRIQDIELKSVTSNNOVSHTHSENDLPHENCM 934
Db 992 IREELGRIBEOHRELEVKQLEQALRIQDIELKSVTSNNOVSHTHSENDLPHENCM 1051
Qy 935 KKEIAMLKLEVATLKHQHVKNKYFEDIKILQEKNAELQMTLKLKQKVTVKRASQYREQ 994
Db 1052 KKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKESLTKRASQYSGQ 1111
Qy 995 LKVLTAENTMLTSKLKE 1011
Db 1112 LKVLIAENTMLTSKLKE 1128

RESULT 11
US-10-007-805-573
; Sequence 573, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-573

Query Match 59,3%; Score 3069; DB 13; Length 1349;
Best Local Similarity 57,3%; Pred. No. 1.8e-152;
Matches 663; Conservative 76; Mismatches 146; Indels 272; Gaps 12;

Qy 1 MVATLLSYGAVIEQNKASLTPLLAIQKRSQTVFELLTKNANANAFNSKCTALMLAI 60
Db 106 VVAKLLSHGAVIEVHNKASLTPLLSITKRSQIVFELLTKNANANAVNKYKCTALMLAV 165
Qy 61 CEGSEISVGMLLQNVDFABDIHGITAERYAAAGVNYIHQQLLEHRLKLPKNPONTNP 120
Db 166 CHGLSEIVGMLLQNVDFVFAADICGVTAERYAVTCGFHIEHQIMEYIRKLSKNHNTNP 225
Qy 121 EGTSTGTPDEAFLA-----ERTPDTAESLLEKTPDEAALVE 158
Db 226 EGTSGTGPDEAFLAERTPDTAESLVEKTPDEAALVETPDTAESLVEKTPDEAALVE 285
```

RESULT 12

US-10-076-622-573

; Sequence 573, Application US/10076222

```
Publication No. US20030023036A1
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: Persing, David H.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C11
CURRENT APPLICATION NUMBER: US/10/076.622
CURRENT FILING DATE: 2002-02-13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 573
LENGTH: 1349
TYPE: PRT
ORGANISM: Homo sapiens
US-10-076-622-573

Query Match 59.3%; Score 3069; DB 14; Length 1349;
Best Local Similarity 57.3%; Pred. No. 1.8e-152;
Matches 663; Conservative 76; Mismatches 146; Indels 272; Gaps 12;

QY 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNSKCTALMLAI 60
Db 106 VVAKLLSHGAVIEVHNKASLTPLLLSITKRSQIVFELLTKNANANAVNKYKCTALMLAV 165
QY 61 CEGSEIVGMLLOQNVDFAEIDHIGITAERYAARGVNYTHOOLLEHRLKPKNPQNTNP 120
Db 166 CHGLSEIVGMLLOQNVDFAAADICGVTAHYAVTCGFGHHIHEQIMEYIRKLSKNHQTNP 225
QY 121 EGTSTGTDPDEAAPLA-----ERTPDATLESLEKTPDAAALIVE 158
Db 226 EGTAGTDPDEAAPLAERTPDATLESVEKTPDDEAAPLVERTPDPAESLVEKTPDDEAASIVE 285
QY 159 GTSAKIOCLGATSGKPFQSTTEPRKILRPTKTSKFSWP----- 200
Db 286 GTSKIOCLERATSGKPFQSAETPREITSKETSFTWPAKGRPRKIAWEKEDTPR 345
QY 201 -----AKERSKITWEKETSVKTECVAGTVPNKTEVLEKGTSMIACP 244
Db 346 EIMSPAKTSEKFTWAAKGRPRKIAWEKETPVKTCGVARTSNKRVLEKGRKMIACP 405
QY 245 TKETSTKASTNVDSVSEPIFLFGRTRIENSQCTKVEEDFNLATKIISAAQNYTCLP 304
Db 406 TKESSTKASAN----- 416
QY 305 DATYOKOIKTINHIEDQMPSESREDEDEYSDWDSGLFESSAKTQVCIPESMYOKVME 364
Db 417 -----DQRFPSKQEEDEYSCDSRLFESSAKIQVCIPESYQKVME 460
QY 365 INREVEELPEKPSAEKPAVEMQKTPNKAFLKNEQTLRAAQMPSESQKQDDENSWS 424
Db 461 INREVEEPKPSAEKPALEMONSPNKAFLKNEQTLRAQMPFPESQKQDYENSWS 520
QY 425 ESPCTVSQKQVYLKPAHQEFDLTSGLKLEESPVKDGLLKPTCGRKVSLPNKALELXDR 484
Db 521 ESLCTVSQKQVCLPKAQHQEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELXDM 580
QY 485 EFTKASPDQGLLKPTCGRKVSLPNKALELKDRETLKASPDNDGLLKPTCGRKVSLPN 544
Db 581 QTFKAEPKPSAFEP-----ATEMOK-----SVFN 606
QY 545 KALELKDRFTKAAQWFPSESQKQDEENSWDFESFLETLNDVCLPKAQHQEFDLTS 604
Db 607 KALELKNEQTLRADETLFSESQKQDYENSWDSTESLCTVSQKQVCLPKAQHQEIDKIN 666
QY 605 GKLEESPDQGLLKPTCGMKISLPNKALELKDRETFKADSVSSVSTFSLFGKPTTENSQ 664
Db 667 GKLEGSFVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPEK-PSAF-----EPAIEMOK 721
QY 665 STKVEEDFNLTTEGATKVTTCQOERDGIIRAPQDQTNKNTSELGKEDTKTSDSE 724
Db 722 SV-----PNKALELKNEQTLRADEILP-----SESKQOYEESWS 759

RESULT 13
US-10-124-805-573
Sequence 573, Application US/10124805
Publication No. US20030166022A1
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: Persing, David H.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C12
CURRENT APPLICATION NUMBER: US/10/124.805
CURRENT FILING DATE: 2002-04-15
NUMBER OF SEQ ID NOS: 627
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 573
LENGTH: 1349
TYPE: PRT
ORGANISM: Homo sapiens
US-10-124-805-573

Query Match 59.3%; Score 3069; DB 14; Length 1349;
Best Local Similarity 57.3%; Pred. No. 1.8e-152;
Matches 663; Conservative 76; Mismatches 146; Indels 272; Gaps 12;

QY 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNSKCTALMLAI 60
Db 106 VVAKLLSHGAVIEVHNKASLTPLLLSITKRSQIVFELLTKNANANAVNKYKCTALMLAV 165
QY 61 CEGSEIVGMLLOQNVDFAEIDHIGITAERYAARGVNYTHOOLLEHRLKPKNPQNTNP 120
Db 166 CHGLSEIVGMLLOQNVDFAAADICGVTAHYAVTCGFGHHIHEQIMEYIRKLSKNHQTNP 225
QY 121 EGTSTGTDPDEAAPLA-----ERTPDATLESLEKTPDAAALIVE 158
Db 226 EGTAGTDPDEAAPLAERTPDATLESVEKTPDDEAAPLVERTPDPAESLVEKTPDDEAASIVE 285
QY 159 GTSAKIOCLGATSGKPFQSTTEPRKILRPTKTSKFSWP----- 200
Db 286 GTSKIOCLERATSGKPFQSAETPREITSKETSFTWPAKGRPRKIAWEKEDTPR 345
QY 201 -----AKERSKITWEKETSVKTECVAGTVPNKTEVLEKGTSMIACP 244
Db 346 EIMSPAKTSEKFTWAAKGRPRKIAWEKETPVKTCGVARTSNKRVLEKGRKMIACP 405
QY 245 TKETSTKASTNVDSVSEPIFLFGRTRIENSQCTKVEEDFNLATKIISAAQNYTCLP 304
Db 406 TKESSTKASAN----- 416
QY 305 DATYOKOIKTINHIEDQMPSESREDEDEYSDWDSGLFESSAKTQVCIPESMYOKVME 364
Db 417 -----DQRFPSKQEEDEYSCDSRLFESSAKIQVCIPESYQKVME 460
QY 365 INREVEELPEKPSAEKPAVEMQKTPNKAFLKNEQTLRAAQMPSESQKQDDENSWS 424
Db 461 INREVEEPKPSAEKPALEMONSPNKAFLKNEQTLRAQMPFPESQKQDYENSWS 520
QY 425 ESPCTVSQKQVYLKPAHQEFDLTSGLKLEESPVKDGLLKPTCGRKVSLPNKALELXDR 484
Db 521 ESLCTVSQKQVCLPKAQHQEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELXDM 580
QY 485 EFTKASPDQGLLKPTCGRKVSLPNKALELKDRETLKASPDNDGLLKPTCGRKVSLPN 544
Db 581 QTFKAEPKPSAFEP-----ATEMOK-----SVFN 606
QY 545 KALELKDRFTKAAQWFPSESQKQDEENSWDFESFLETLNDVCLPKAQHQEFDLTS 604
Db 607 KALELKNEQTLRADETLFSESQKQDYENSWDSTESLCTVSQKQVCLPKAQHQEIDKIN 666
QY 605 GKLEESPDQGLLKPTCGMKISLPNKALELKDRETFKADSVSSVSTFSLFGKPTTENSQ 664
Db 667 GKLEGSFVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPEK-PSAF-----EPAIEMOK 721
QY 665 STKVEEDFNLTTEGATKVTTCQOERDGIIRAPQDQTNKNTSELGKEDTKTSDSE 724
Db 722 SV-----PNKALELKNEQTLRADEILP-----SESKQOYEESWS 759
```

```

Db 346 EIMGPAKETSBEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACP 405
Qy 245 TKETSTKASTNDVSSVEPIFSLFQTRTIENSQCTKVEEDFNLATKIISKAAQNYTCLP 304
Db 406 TKESSTKASAN----- 416
Qy 305 DATYQDKITINHIEDQMPSESKREDEBYSDSGSLPSSAKTQVCIPESMYQKWE 364
Db 417 -----DQFPSESKQDEDEYSCDSRSLFESSAKIQVCIPESYQKWE 460
Qy 365 INREVEELPEKPSAFKPAVENQKTPVNPKAFLKNEOTLRAAQMPSSSKQKDDSENSWDS 424
Db 461 INREVEEPKPSAFKPAIEMQNSVPNKAFLKNEOTLRAQMPSSSKQKDDSENSWDS 520
Qy 425 ESPCETVSQKDVYLPKATHQKEFDLTSGKLEESPVKDGILLKPTCGRKVSLPNKALELKD 484
Db 521 ESLCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELKD 580
Qy 485 ETFKAESPDKGILLKPTCGRKVSLPNKALELKDRETLKAESPNDGILLKPTCGRKVSLPN 544
Db 581 QTFKAESPDKGILLKPTCGRKVSLPNKALELKDRETLKAESPNDGILLKPTCGRKVSLPN 606
Qy 545 KALELKORETFKAAQMPSESKQKDDSENSWDSFLETLONDVCLPKATHQKEFDLTSL 604
Db 607 KALELKNEQTRADEILPESKQKDDSENSWDSFLETLONDVCLPKATHQKEFDLTSL 666
Qy 605 KGLSESPDKGILLKPTCGMKISLNPNALELKDRETFKAEDEVSSVESTPSLFGKPTTENSQ 664
Db 667 KGLSESPDKGILLKPTCGMKISLNPNALELKDRETFKAEDEVSSVESTPSLFGKPTTENSQ 721
Qy 665 STKVEEDFNLTKEGATKVTGQOERDIGIIERAPQDQTNKMPSELGRKEDTKSTDSSE 724
Db 722 SV-----PNKALELKDRETFKAEDEVSSVESTPSLFGKPTTENSQ 759
Qy 725 IISVSDTQNYECLPEATYQKEIKTTNGKIEES----- 756
Db 760 SLCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMD 819
Qy 757 -----PEKPSHPEPATQNSVPNKGLEWKNQOTLRA----- 788
Db 820 TFKAEPEKPSAFKPAIEMQNSVPNKALELKDRETLKAESPNDGILLKPTCGRKVSLPN 879
Qy 789 -----DSTTLISKILDALPSCERGRELKKNCEQITA 819
Db 880 LRETVSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDVHSCERARELQKHCEQRTG 939
Qy 820 KMEQKXKFCVLOKELSEAKIKSOLENQKAKWQELCSVRLPNQBEERKRNVDILKEK 879
Db 940 KMEQKXKFCVLOKELSEAKIKSOLENQKAKWQELCSVRLPNQBEERKRNVDILKEK 999
Qy 880 IRP-----EQLRKKLEVKHOLEOTLRIOQIELKSVTSNLNQVSHTHSENDLPHENCM 934
Db 1000 IRELGRIEEOHRELEVKQLEQALRIQDIELKSVESNLNQVSHTHSENLPHENCM 1059
Qy 935 KKEIAMLKLEATLKHQHVKNKYFEDIKIOEKNAELQMTLKLQKVTTKRASQYREQ 994
Db 1060 KKEIAMLKLEATLKHQHVKNKYFEDIKILKEKNAELQMTLKLKESLTKRASQYREQ 1119
Qy 995 LKVTAENTMLTSKLKE 1011
Db 1120 LKVIAENTMLTSKLKE 1136

```

RESULT 14

```

US-10-007-805-553
; Sequence 553, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.

```

```

; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 553
; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-553

```

```

Query Match 45.9%; Score 2374; DB 13; Length 1013;
Best Local Similarity 56.0%; Pred. No. 3.4e-116;
Matches 512; Conservative 61; Mismatches 126; Indels 216; Gaps 9;

Qy 187 LRPTKTSKPSWPAKERSRKITWEEKETSVKTECVAGVTNKTVELEKGTSMIACPTK 246
Db 12 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACP 71
Qy 247 ETSKASTNDVSSVEPIFSLFQTRTIENSQCTKVEEDFNLATKIISKAAQNYTCLPDA 306
Db 72 ESTKASAN----- 80
Qy 307 TYQDKITINHIEDQMPSESKREDEBYSDSGSLPSSAKTQVCIPESMYQKWEIN 366
Db 81 -----DQFPSESKQDEDEYSCDSRSLFESSAKIQVCIPESYQKWEIN 126
Qy 367 REVEELPEKPSAFKPAVENQKTPVNPKAFLKNEOTLRAAQMPSSSKQKDDSENSWDS 426
Db 127 REVEEPKPSAFKPAIEMQNSVPNKAFLKNEOTLRAQMPSSSKQKDDSENSWDS 186
Qy 427 PCETVSQKDVYLPKATHQKEFDLTSGKLEESPVKDGILLKPTCGRKVSLPNKALELKD 486
Db 187 LCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELKD 246
Qy 487 FKABSPDKGILLKPTCGRKVSLPNKALELKDRETLKAESPNDGILLKPTCGRKVSLPNKA 546
Db 247 FKABSPDKGILLKPTCGRKVSLPNKALELKDRETLKAESPNDGILLKPTCGRKVSLPNKA 284
Qy 547 LELKDRETFKAAQMPSSSKQKDDSENSWDSFLETLONDVCLPKATHQKEFDLTSLGK 606
Db 285 -----EILPSESKQKDEYSDSWDESLETCETVSQKDVCLPKAAHQKEIDKINGK 332
Qy 607 LEESPDGILLKPTCGMKISLNPNALELKDRETFKAEDEVSSVESTPSLFGKPTTENSOST 666
Db 333 LEGSPGKDGILLKANGCMKVSIPKALELMDQTFKABPEK-PSAF-----FPAIEMQNSV 387
Qy 667 KVEEDFNLTKEGATKVTGQOERDIGIIERAPQDQTNKMPSELGRKEDTKSTDSBII 726
Db 388 -----PNKALELKDRETFKAEDEVSSVESTPSLFGKPTTENSOST 425
Qy 727 SVSDTQNYECLPEATYQKEIKTTNGKIEES----- 756
Db 426 CETVSQKDVCLPKAAHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELMDQTF 485
Qy 757 -----PEKPSHPEPATQNSVPNKGLEWKNQOTLRA----- 788
Db 486 KAEPEKPSAFKPAIEMQNSVPNKALELKDRETFKAEDEVSSVESTPSLFGKPTTENSOST 545
Qy 789 -----DSTTLISKILDALPSCERGRELKKNCEQITA 821
Db 546 ETVSQKDVCPKATHQKEMDKISGKLEEDSTSLSKILDVHSCERARELQKHCEQRTG 605
Qy 822 EQMKXKFCVLOKELSEAKIKSOLENQKAKWQELCSVRLPNQBEERKRNVDILKEKIR 881

```

Db 606 EQMKKFCVLRKKLSEAKEIKSLENOQKVKWEQELCSVRLTLNQEBEKRNDILNEKIR 665

Qy 882 P-----EQLRKKLEVKHQLEQTLRIQDIELKSVTSNLSNVSHTHSENDLFHENCMLKK 936

Db 666 EELGRIEEQRKELEVKQQLAQALRIQDIELKSVESNLSNVSHTHSENYLLHENCMLKK 725

Qy 937 EIAMLKLEVATLKHQHVKENKYPEDIKILQEKNAELQMTLKLKQKTVTKRASQYREOLK 996

Db 726 EIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKBEESLTKRASQYSGQLK 785

Qy 997 VLTAENTMLTSKLKE 1011

Db 786 VLIAENTMLTSKLKE 800

RESULT 15

US-10-076-622-553

; Sequence 553, Application US/10076622

; Publication No. US20030023036A1

; GENERAL INFORMATION:

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Persing, David H.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C11

; CURRENT APPLICATION NUMBER: US/10/076,622

; CURRENT FILING DATE: 2002-02-13

; NUMBER OF SEQ ID NOS: 627

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 553

; LENGTH: 1013

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-076-622-553

Query Match 45.9%; Score 2374; DB 14; Length 1013;

Best Local Similarity 56.0%; Pred No. 3 4e-116;

Matches 512; Conservative 61; Mismatches 126; Indels 216; Gaps 9;

Qy 187 LRPTKETSEKFSWPAKERSKITWEEKTSVKTECVAGVTPNKTVELEKGTSGNMIACPTK 246

Db 12 MSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTGCVARVTSNKTVKLEKRSKMIACPTK 71

Qy 247 EISTKASTNVDSVSEPIESLFGTRTIENSQCTKVEEDFNLATKIISKSAQNYTCLPDA 306

Db 72 ESSTKASAN-----80

Qy 307 TYQKDIKTINHKIEDQMPFSESKREDEEYSWDSGLPSSAKTQVCIPESMYQKWEIN 366

Db 81 -----DQRFSESKQEBDEEYSCDSRSLPSSAKIQVCIPESYQKWEIN 126

Qy 367 REVEELPEKPSAPKPAVENQKTVPNKAFELKNEQTLRAAQMPFSESKQKQDDENSWDS 426

Db 127 REVEEPKPSAPKPAIEMQNSVPNKAFELKNEQTLRADPMPFSESKQKQDDENSWDS 186

Qy 427 PCETVSKQVYLPKATHQKEFDTLGKLEESPVKQGLLKPTCGRKVSLPNKALELKDR 486

Db 187 LCETVSKQVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELKDQMT 246

Qy 487 FKAESPDKQGLLKPTCGRKVSLPNKALELKDRTELKAEPSDNDGLLKPTCGRKVSLPNKA 546

Db 247 FKAEPKPSAPFAPATEMOKSVPNKALELKNEQTLRAD-----284

Qy 547 LELKDRETFKAQOMFSESKQDDENSWDFSEFLETLLONDVCLPKATHQKEFDTLGK 606

Db 285 -----EILPSESKQDYENSNDTESLCEVTVSQKQVCLPKAAHQKEIDKINGK 332

Qy 607 LEESPDKQGLLKPTCGMKKISLPNKALELKDRETFKAEDVSSVBSTFSLFGKPTTENSQST 666

Db 333 LEGSPGKDGILLKANCCKMKVSIPTKALELMDQMTFKAEPPEK-PSAF-----EPAIEMOKSV 387

Qy 667 KVEEDFNLTTKEGATKTVTGQQRDGIIEERAFQDQTNKMPSELGKEDTKTSDSSEII 726

Db 388 -----PNKALELKNEQTLRADEILP-----SESKQDYESSWDS 425

Qy 727 SVSDTONYECLEPATYQKEIKTTNGKIES-----756

Db 426 CETVSQKQVCLPKAAHQKEIDKINGKLEESPNQDGLKSPCRMKVSIPTKALELMDQMTF 485

Qy 757 ----PEKPSHPBPATEMOMNSVPNKGLWKNKQTLRA-----788

Db 486 KAEPPKPSAPFAPATEMOKSVPNKALELKNEQTLRADQMPFSESKQKQVENSNDSESLR 545

Qy 789 -----DSTTLSKILDALPSCERRELKKNCEQITAKM 821

Db 546 ETVSQKQVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTGKM 605

Qy 822 EQMKKFCVLRKKLSEAKEIKSLENOQKVKWEQELCSVRLTLNQEBEKRNDILNEKIR 881

Db 606 EQMKKFCVLRKKLSEAKEIKSLENOQKVKWEQELCSVRLTLNQEBEKRNDILNEKIR 665

Qy 882 P-----EQLRKKLEVKHQLEQTLRIQDIELKSVTSNLSNVSHTHSENDLFHENCMLKK 936

Db 666 EELGRIEEQRKELEVKQQLAQALRIQDIELKSVESNLSNVSHTHSENYLLHENCMLKK 725

Qy 937 EIAMLKLEVATLKHQHVKENKYPEDIKILQEKNAELQMTLKLKQKTVTKRASQYREOLK 996

Db 726 EIAMLKLEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKBEESLTKRASQYSGQLK 785

Qy 997 VLTAENTMLTSKLKE 1011

Db 786 VLIAENTMLTSKLKE 800

Search completed: February 19, 2005, 02:33:53
Job time : 101.958 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:57:06 ; Search time 27.5342 Seconds
(without alignments)
2740.961 Million cell updates/sec

Title: US-09-602-362E-27

Perfect score: 5173

Sequence: 1 MVATLLSYGAVIEVQNKSL.....REQLKVLTAENTMLTSLKE 1011

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2372.5	45.9	1095	4	US-09-620-405B-493
2	2372.5	45.9	1095	4	US-09-834-759-493
3	2367	45.8	1002	4	US-09-620-405B-475
4	2367	45.8	1002	4	US-09-604-287A-475
5	2367	45.8	1002	4	US-09-834-759-475
6	2367	45.8	1002	4	US-09-590-751A-475
7	2367	45.8	1002	4	US-09-551-621-475
8	1756.5	34.0	743	4	US-09-620-405B-494
9	1756.5	34.0	743	4	US-09-834-759-494
10	1751	33.8	650	4	US-09-620-405B-469
11	1751	33.8	650	4	US-09-433-826B-469
12	1751	33.8	650	4	US-09-604-287A-469
13	1751	33.8	650	4	US-09-834-759-469
14	1751	33.8	650	4	US-09-590-751A-469
15	1751	33.8	650	4	US-09-551-621-469
16	1542	29.8	512	4	US-09-451-739H-16
17	1414.5	27.3	445	4	US-09-620-405B-473
18	1414.5	27.3	445	4	US-09-433-826B-473
19	1414.5	27.3	445	4	US-09-604-287A-473
20	1414.5	27.3	445	4	US-09-834-759-473
21	1414.5	27.3	445	4	US-09-590-751A-473
22	1414.5	27.3	445	4	US-09-551-621-473
23	1384	26.8	466	4	US-09-620-405B-472
24	1384	26.8	466	4	US-09-433-826B-472
25	1384	26.8	466	4	US-09-604-287A-472
26	1384	26.8	466	4	US-09-834-759-472
27	1384	26.8	466	4	US-09-590-751A-472

28	1384	26.8	466	4	US-09-551-621-472	Sequence 472, App
29	884.5	17.1	432	4	US-09-389-681-181	Sequence 181, App
30	884.5	17.1	432	4	US-09-620-405B-181	Sequence 181, App
31	884.5	17.1	432	4	US-09-339-338-181	Sequence 181, App
32	884.5	17.1	432	4	US-09-433-826B-181	Sequence 181, App
33	884.5	17.1	432	4	US-09-604-287A-181	Sequence 181, App
34	884.5	17.1	432	4	US-09-285-480-181	Sequence 181, App
35	884.5	17.1	432	4	US-09-834-759-181	Sequence 181, App
36	884.5	17.1	432	4	US-09-590-751A-181	Sequence 181, App
37	884.5	17.1	432	4	US-09-551-621-181	Sequence 181, App
38	798.5	15.4	228	4	US-09-620-405B-470	Sequence 470, App
39	798.5	15.4	228	4	US-09-433-826B-470	Sequence 470, App
40	798.5	15.4	228	4	US-09-604-287A-470	Sequence 470, App
41	798.5	15.4	228	4	US-09-834-759-470	Sequence 470, App
42	798.5	15.4	228	4	US-09-590-751A-470	Sequence 470, App
43	798.5	15.4	228	4	US-09-551-621-470	Sequence 470, App
44	481	9.3	179	4	US-09-389-681-179	Sequence 179, App
45	481	9.3	179	4	US-09-620-405B-179	Sequence 179, App

ALIGNMENTS

RESULT 1

US-09-620-405B-493

; Sequence 493, Application US/09620405B

; Patent No. 6528054

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yugui

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Hepler, William T.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C8

; CURRENT APPLICATION NUMBER: US/09/620.405B

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 495

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 493

; LENGTH: 1095

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: variant

; LOCATION: (1)...(1095)

; OTHER INFORMATION: Xaa = Any amino acid

US-09-620-405B-493

Query Match 45.9%; Score 2372.5; DB 4; Length 1095;

Best Local Similarity 52.2%; Pred. No. 1.9e-156; Indels 261; Gaps 13;

Matches 52; Conservative 73; Mismatches 131;

Qy 104 LLEHRLKLPKNQNPETGTGTPDEAALPAERTPTAESLLKTPDEAARLVETGSAK 163

Db 24 LLENVSKTINPQVSKTE-----YKELLQEFIDDNA-----TTNA 58

Qy 164 IQCLGKATSKFPOSTETPRKI-----LAPTKEKFSWPAKERSRK 207

Db 59 IDELKEC-----FLNQTDETLSNVFVFMQLIYDSSLCDLFMSPAKETSEKFTWAAKGRPK 114

Qy 208 ITWEKETSVKTECVAGVTNKTVEKGTSMNIACPTKETSTKASTNVDSVVEPIFSL 267

Db 115 IAWBKETPVKTGCVARVTSNKTVKLEGRSKMTACPTKESSTKASAN-----162

Qy 268 FGTRTIENSQCTKVEEDFNLATKIISKSAQNTCLPDATYQKDIKTINHKEIDQMPFSE 327

Db 163 -----DQRFPSF 169

Qy 328 SKREDEYSWDSGLFESSAKTQVCIPESMYQKMEINREVELPEKPSAFKAVENQK 387

Db 163 -----DQRFPSF 169

170 SKQEDBEYSCDSRSLFESSAKIQVCIPESYQKVMINREVEEPPKPSAFKPAIEMQN 229
388 TVPNKAPFELKNEQTLRAAQMPPSSKQKDBEENSWSDESPCETVSQKDVYLPKATHQKEP 447
230 SVPNKAPFELKNEQTLRADPMFPSPSKQKDYENSWSDESLCETVSQKDVCLPKATHQKEI 289
448 DTLGKLEESPVKQGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 507
290 DKINGKLEESPNKQGLLKATCMKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 349
508 LPNKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 567
350 VENKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 375
568 KDBEENSWSDESLCETVSQKDVCLPKATHQKEPDLTGKLEESPDGKGLLKPTCGRKVS 627
376 KDBEENSWSDESLCETVSQKDVCLPKATHQKEIDKINGKLEESPDGKGLLKPTCGRKVS 435
628 PNKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 687
436 PTKALELMDQTFKAEPPEK-PSAF-----EPAIEMQKSV-----PNKALELK 477
688 QERDIGIIRAPQDQTNMPTSELGRKEDTKSTDSSEIISVSDTQNYECLPEATYQKEIK 747
478 NEQTLRADEILP-----SESKQDYESSWSDESLCETVSQKDVCLPKATHQKEID 528
748 TTNGKIEES-----EPAIEMQKSV-----EILPSESKQ 773
529 KINGKLEESPDGGLLKATCMKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 588
774 PNKGLEWKNKQTLRA-----EPAIEMQKSV-----EILPSESKQ 788
589 PNKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 648
789 -----DSTTLGKLEESPDGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 789
649 ISGKLEESPDGGLLKATCMKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 708
843 SOLENOKAKWEOLCSVRLPLNQBEERKNVDILKKEIRP-----EOLRKLKLEVKHOLE 897
709 SOLENOKAKWEOLCSVRLPLNQBEERKNVDILKKEIRP-----EOLRKLKLEVKHOLE 768
898 QTLRIODIELKSVTSLNQVSHTHSENDLFHENCMLKKEIAMLKLEIATLKHQVQKEN 957
769 QALRIODIELKSVTSLNQVSHTHSENDLFHENCMLKKEIAMLKLEIATLKHQVQKEN 828
958 KYFEDIKILKKEKNAELQMTLKLKQKVTTRASQYREQLKVLTAENTMLTSKLKE 1011
829 KYFEDIKILKKEKNAELQMTLKLKQKVTTRASQYREQLKVLTAENTMLTSKLKE 882

RESULT 2
US-09-834-759-493
Sequence 493, Application US/09834759
Patent No. 6680197
GENERAL INFORMATION:
APPLICANT: Jiang, Yugui
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.470C9
CURRENT FILING DATE: 2001-04-13
NUMBER OF SEQ ID NOS: 547
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 493
LENGTH: 1095
TYPE: PRT

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: variant
LOCATION: (1)...(1095)
OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493
Query Match 45.9%; Score 2372.5; DB 4; Length 1095;
Best Local Similarity 52.2%; Pred. No. 1.9e-156;
Matches 529; Conservative 73; Mismatches 151; Indels 261; Gaps 13;
104 LLEHRIKLPKNQNTNPEGTSTGTGTPDEAALPABRTPDTAESLLEKTPDEAALVSGTSAK 163
24 LLENVISTKINQVSKTE-----YKELLQBFIDNA-----TTNA 58
164 IQCLGKATSGKFEQSTETPRKI-----LRPTKETSFKFWPAKERSRK 207
59 IDELKEC-----FVNQIDETLSNVFVPMQIYDSSILCDLPMSPAKETSEKFTWAAGRPRK 114
208 ITWEEKETSVKTECVAGVTPNKTVELEKTSNMIACTPKETSTKASTNVDSVVEPISL 267
115 IAWKKEKTPVKTGCVARVTSNKTVLEKGRSMIACTPKETSTKASTNVDSVVEPISL 162
268 FGTRTISNQCQTVKEDFNLAIKIISKSAQNYTCLPDATYQKDIKTNHKTEDQMPFSE 327
163 -----DQFPFSE 169
328 SKREDEBEYSWSGSLFESSAKTQVCIPESMYQKVMINREVEEPPKPSAFKPAIEMQK 387
170 SKQEDBEYSCDSRSLFESSAKIQVCIPESYQKVMINREVEEPPKPSAFKPAIEMQN 229
388 TVPNKAPFELKNEQTLRAAQMPPSSKQKDBEENSWSDESPCETVSQKDVYLPKATHQKEP 447
230 SVPNKAPFELKNEQTLRADPMFPSPSKQKDYENSWSDESLCETVSQKDVCLPKATHQKEI 289
448 DTLGKLEESPVKQGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 507
290 DKINGKLEESPNKQGLLKATCMKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 349
508 LPNKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 567
350 VENKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 375
568 KDBEENSWSDESLCETVSQKDVCLPKATHQKEPDLTGKLEESPDGKGLLKPTCGRKVS 627
376 KDBEENSWSDESLCETVSQKDVCLPKATHQKEIDKINGKLEESPDGKGLLKPTCGRKVS 435
628 PNKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 687
436 PTKALELMDQTFKAEPPEK-PSAF-----EPAIEMQKSV-----PNKALELK 477
688 QERDIGIIRAPQDQTNMPTSELGRKEDTKSTDSSEIISVSDTQNYECLPEATYQKEIK 747
478 NEQTLRADEILP-----SESKQDYESSWSDESLCETVSQKDVCLPKATHQKEID 528
748 TTNGKIEES-----EPAIEMQKSV-----EILPSESKQ 773
529 KINGKLEESPDGGLLKATCMKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 588
774 PNKGLEWKNKQTLRA-----EPAIEMQKSV-----EILPSESKQ 788
589 PNKALELKDRETFKAESPDKGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 648
789 -----DSTTLGKLEESPDGGLLKPTCGRKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 789
649 ISGKLEESPDGGLLKATCMKVSIPNKALELKDRETFKAESPDKGGLLKPTCGRKVS 708
843 SOLENOKAKWEOLCSVRLPLNQBEERKNVDILKKEIRP-----EOLRKLKLEVKHOLE 897
709 SOLENOKAKWEOLCSVRLPLNQBEERKNVDILKKEIRP-----EOLRKLKLEVKHOLE 768
898 QTLRIODIELKSVTSLNQVSHTHSENDLFHENCMLKKEIAMLKLEIATLKHQVQKEN 957


```

; NAME: ALAN
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-475

```

Query Match	45.8%;	Score 2367;	DB 4;	Length 1002;
Best Local Similarity	55.8%;	Pred. No. 4.1e-150;		
Matches 511; Conservative 62; Mismatches 126; Indels 216; Gaps 9				
Qy	187	LRPKYTSKFSWPAKRSRKLTWEEKTSVKTECVAGVTNKTVELSGKTSNMIACTPK	246	
	:	:	:	
Db	1	MSPAKETSEKPTAAAGCPRIAWKEKETPVKTGCVARVTSNKTVELSGRSKMIACTPK	60	
	:	:	:	
Qy	247	ETSTKASTNDVSSVEPIFLSGTRTENSQCTKVEEDFNLATKILSKSAAQNYTCLPDA	306	
	:	:	:	
Db	61	ESSTKASAN-----	69	
Qy	307	TYQDKITINHKIEDOMPPSESKREEDSEYSDSGSLPSSAKTOVCIPESMYQKWEIN	366	
	:	:	:	

```
Db 70 -----DQFPSEKQDEDEYSCDSRSLFESSAKIQVCIPESYQKVMWEIN 115
Qy 367 REVEELPKPSAFKPAVEMQKTVNPKAFELKNEQTLRAAQMPFSESKQKDEENSWDES 426
Db 116 REVEEPKPSAFKPAIEMQNSVNPKNAPKELKNEQTLRADPMFPFSESKQKDEENSWDES 175
Qy 427 PCETVSQKDVLPKATHQKEEDTLGKLEESPVKDGKLLKPTCGRKVSLPNKALELKDRET 486
Db 176 LCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGKLLKATCGMKVSIPTKALELKDRET 235
Qy 487 FKAESPDKDGLLKPTCGRKVSLPNKALELKDRETILKAESEPNKDGKLLKPTCGRKVSLPNKA 546
Db 236 FKAEPKPSAFKPAIEMQNSVNPKNAPKELKNEQTLRAD- 273
Qy 547 LELKDRETFAAQMPFSESKQKDEENSWDESLETLQNDVCLPKATHQKEEDTLGK 606
Db 274 -----EILPSESKQKDEESWDESLETLQNDVCLPKATHQKEIDKINGK 321
Qy 607 LEESPDKDGKLLKPTCGMKISLPNKALELKDRETFAEDVSSVESTFSLFGKPTTENSOST 666
Db 322 LEGSPVKDGLLKANGCMKVSIPKALELMDQMTFKAEPPEK-PSAF-----EPAIEMQKSV 376
Qy 667 KVEDFNLTKEGATKVTGQERDIIIRAPQDQTNKMPSELGRKEDTKSDSII 726
Db 377 -----PNKALELNEQTLRADEILP-----SESKQKDEESWDESIL 414
Qy 727 SVSDTQNYCELPKATYQKEIKTTNGKIEES 756
Db 415 CETVSQKDVCLPKAXHQKEIDKINGKLEESPNKDGKLLKAPCRMKVSIPKALELMDQMTF 474
Qy 757 -----PEKPSHPEPATMONGVNPKNKLEWKNQTLRA- 788
Db 475 KAEPPEKPSAFKPAIEMQNSVNPKNAPKELKNEQTLRADQMPFSESKQKXVEENSWDESILR 534
Qy 789 -----DSTTLTKILDALPSCERGLKKNQCEITAKM 821
Db 535 ETVSQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGM 594
Qy 822 EQMKKFCVLQKELSEAKEIKSQLENQAKWEQELCSVRLPLNQVEEKRRNVNVLKKEIR 881
Db 595 EQMKKFCVLKKELESEAKEIKSQLENQAKWEQELCSVRLPLNQVEEKRRNADILNEKIR 654
Qy 882 P-----EEQLRKKLEVKHOLEQTLRIQDIELKSVTSNINQVSHTHESNDLPHENCMLKK 936
Db 937 EIAMLKLEVATLKHQVQKVENKYFEDIKILQKNAELQMTLKLKOKTVTKASQYREOLK 996
Db 715 EIAMLKLEIATLKHQVQKVENKYFEDIKILKKEKNAELQMTLKLKBEESLTKRASQYSGOLK 774
Qy 997 VITAENTMLTSKLKE 1011
Db 775 VLIAENTMLTSKLKE 789
```

RESULT 5

```
US-09-834-759-475
; Sequence 475, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
; US-09-834-759-475
```

```
Query Match 45.8%; Score 2367; DB 4; Length 1002;
Best Local Similarity 55.8%; Pred. No. 4.1e-156;
Matches 511; Conservative 62; Mismatches 126; Indels 216; Gaps 9;

Qy 187 LRPTKETSEKPSAPKERSRKITWEKETSVKTECVAGVTPNKTEVLEKGTNNIACTPK 246
Db 1 MSPAKETSEKTPWAAGRPKRIAWERKETPVKGCVARVTSNKTVKLEGRSRMIACTPK 60
Qy 247 ETSTKASTNVDSVSEPIFSLFGTRTTIENSQCTKVEEDFNLATKIISKSAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69
Qy 307 TYQKDIKTIHKKIEDQMPFSESKREDEEYVSWDGSLSFESSAKTQVCIPESMYQKVMWEIN 366
Db 70 -----DQFPSEKQDEDEYSCDSRSLFESSAKIQVCIPESYQKVMWEIN 115
Qy 367 REVEELPKPSAFKPAVEMQKTVNPKAFELKNEQTLRAAQMPFSESKQKDEENSWDES 426
Db 116 REVEEPKPSAFKPAIEMQNSVNPKNAPKELKNEQTLRADPMFPFSESKQKDEENSWDES 175
Qy 427 PCETVSQKDVLPKATHQKEEDTLGKLEESPVKDGKLLKPTCGRKVSLPNKALELKDRET 486
Db 176 LCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGKLLKATCGMKVSIPTKALELKDRET 235
Qy 487 FKAESPDKDGLLKPTCGRKVSLPNKALELKDRETILKAESEPNKDGKLLKPTCGRKVSLPNKA 546
Db 236 FKAEPKPSAFKPAIEMQNSVNPKNAPKELKNEQTLRAD- 273
Qy 547 LELKDRETFAAQMPFSESKQKDEENSWDESLETLQNDVCLPKATHQKEEDTLGSK 606
Db 274 -----EILPSESKQKDEESWDESLETLQNDVCLPKATHQKEIDKINGK 321
Qy 607 LEESPDKDGKLLKPTCGMKISLPNKALELKDRETFAEDVSSVESTFSLFGKPTTENSOST 666
Db 322 LEGSPVKDGLLKANGCMKVSIPKALELMDQMTFKAEPPEK-PSAF-----EPAIEMQKSV 376
Qy 667 KVEDFNLTKEGATKVTGQERDIIIRAPQDQTNKMPSELGRKEDTKSDSII 726
Db 377 -----PNKALELNEQTLRADEILP-----SESKQKDEESWDESIL 414
Qy 727 SVSDTQNYCELPKATYQKEIKTTNGKIEES 756
Db 415 CETVSQKDVCLPKAXHQKEIDKINGKLEESPNKDGKLLKAPCRMKVSIPKALELMDQMTF 474
Qy 757 -----PEKPSHPEPATMONGVNPKNKLEWKNQTLRA- 788
Db 475 KAEPPEKPSAFKPAIEMQNSVNPKNAPKELKNEQTLRADQMPFSESKQKXVEENSWDESILR 534
Qy 789 -----DSTTLTKILDALPSCERGLKKNQCEITAKM 821
Db 535 ETVSQKDVCPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGM 594
Qy 822 EQMKKFCVLQKELSEAKEIKSQLENQAKWEQELCSVRLPLNQVEEKRRNVNVLKKEIR 881
Db 595 EQMKKFCVLKKELESEAKEIKSQLENQAKWEQELCSVRLPLNQVEEKRRNADILNEKIR 654
Qy 882 P-----EEQLRKKLEVKHOLEQTLRIQDIELKSVTSNINQVSHTHESNDLPHENCMLKK 936
Db 937 EIAMLKLEVATLKHQVQKVENKYFEDIKILQKNAELQMTLKLKOKTVTKASQYREOLK 996
Db 715 EIAMLKLEIATLKHQVQKVENKYFEDIKILKKEKNAELQMTLKLKBEESLTKRASQYSGOLK 774
Qy 997 VITAENTMLTSKLKE 1011
Db 775 VLIAENTMLTSKLKE 789
```



```
QY 367 REVELPEKPSAFKPAVEMQKTPVNFPAFELKNEQTLRAAQMPFSESQKQDEENSWSES 426
Db 116 REVEPPKPKPSAFKPAVEMQKTPVNFPAFELKNEQTLRAQMPFSESQKQDEENSWSES 175
QY 427 PCEIVSQDVLPKATHQKEFTLSGKLEESPVKDGLLKPTCGRKVSLPNKALELKDBET 486
Db 176 LCETVSKDVCLPKATHQKEFTLSGKLEESPVKDGLLKPTCGRKVSLPNKALELKDBET 235
QY 487 FKAESPDKDGLLKPTCGRKVSLPNKALELKDBETLSGKLEESPVKDGLLKPTCGRKVSLPNKA 546
Db 236 FKAESPDKDGLLKPTCGRKVSLPNKALELKDBETLSGKLEESPVKDGLLKPTCGRKVSLPNKA 273
QY 547 LELKDRFTKAAQMPFSESQKQDEENSWDFESFLETLQNDVCLPKATHQKEFTLSGK 606
Db 274 -----EILPSSSKQDYESSWDSSELCTVSKQDVCLPKAXHQKEIDKINGK 321
QY 607 LEESPDGKLLKPTCGRKVSLPNKALELKDBETLSGKLEESPVKDGLLKPTCGRKVSLPNKA 666
Db 322 LEGSPKDGKLLKPTCGRKVSLPNKALELKDBETLSGKLEESPVKDGLLKPTCGRKVSLPNKA 376
QY 667 KVEEDFNLTKEGAPKTVTGQERDGIIRAPQDQTNKMPSELGRKEDTKSTSDSII 726
Db 377 -----PNKALELKNEQTLRADEILP-----SESKQDYESSWDSSEL 414
QY 727 SVSDTQNYECLPEATYQKEIKTTNGKIBES----- 756
Db 415 CETVSKQDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTF 474
QY 757 -----PEKSHFEPATEMNSQVFNKLEWKNQOTLRA----- 788
Db 475 KAEPPEKPSAFEPATEMNSQVFNKLEWKNQOTLRA----- 788
QY 789 -----DSTTLKSLDALPSCERGERELKDNCEQITAKM 821
Db 535 ETVSKQDVCPKATHQKEMDKISGKLESTLSKILDTVHSCERARELOKHCEQORTGM 594
QY 822 EQMKNKFCVLOKELSEAKIQLBNQAKNEQELCSVRLPLNQBEERNNVDILKEKIR 881
Db 595 EQMKNKFCVLOKELSEAKIQLBNQAKNEQELCSVRLPLNQBEERNNVDILKEKIR 654
QY 882 P-----EQRLKRLVKGLEQTLRIQDIELKSVTSNLQVSHTHESNDILFHCNMLKK 936
Db 655 BELGRIEBQHRKEVKGLEQTLRIQDIELKSVTSNLQVSHTHESNDILFHCNMLKK 714
QY 937 ETAMLKLEATLKHQVKNKYFEDIKILOKNAELQWTLKIKOKTVTKRASQVREOLK 996
Db 715 ETAMLKLEATLKHQVKNKYFEDIKILOKNAELQWTLKIKOKTVTKRASQVREOLK 774
QY 997 VLTAENTMLTSKIKE 1011
Db 775 VLTAENTMLTSKIKE 789
```

RESULT 8

US-09-620-405B-494

; Sequence 494, Application US/09620405B

; Patent No. 6528054

; GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Hepler, William T.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C8

; CURRENT APPLICATION NUMBER: US/09/620,405B

; CURRENT FILING DATE: 2000-07-20

; NUMBER OF SEQ ID NOS: 495

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 494

; LENGTH: 743

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(743)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-494
```

Query Match 34.0%; Score 1756.5; DB 4; Length 743;

Best Local Similarity 46.0%; Pred. No. 7.8e-114;

Matches 404; Conservative 67; Mismatches 140; Indels 267; Gaps 13;

```
QY 104 LLEHRLKPNKPNQNTPEGTSTGTTPDEAAPLAERTPTDAESLLEKTPDEAARLVEGTSK 163
Db 24 LLENVISKTINPQVSKTE-----YKELLOEFDDNA-----TTNA 58
```

```
QY 164 IQCLGKATSGKFEOSTETPRKI-----LRPTKETSEKTSWPAKERSRK 207
Db 59 IDELKEC-----FLNQTDITLSNVFVFMQLIYDSSLCOLFMSPAKETSEKFTWAAKGRPK 114
```

```
QY 208 ITWEKETSVKTECVAGVTPNKTVELEKGTISNMIACTKETSTKASTNVDVSSVEPIFSL 267
Db 115 IANEKKEPTVKTGCVARVTSNKTVELEKGTISNMIACTKETSTKASTNVDVSSVEPIFSL 162
```

```
QY 268 FGTRTIENSQCTKVEEDFNLATKIIISKAAQNYTCLPDATYQDKIKNHKBIDQMPFSE 327
Db 163 -----DQRFPE 169
```

```
QY 328 SKREDEYSWDSGLSFESSAKTQVCIPESMYQKVINREVELPKPSAFKPAVEMOK 387
Db 170 SKQEBEYSCDSRLSFESSAKIQVCIPESYQKVINREVEPPKPSAFKPAVEMON 229
```

```
QY 388 TVPNKAFELKNEQTLRAAQMPFSESQKQDEENSWDSSELCTVSKQDVCLPKATHQKEF 447
Db 230 SVPNKAFELKNEQTLRAQMPFSESQKQDYENSWDSSELCTVSKQDVCLPKATHQKEI 289
```

```
QY 448 DTLGKLEESPVKDGKLLKPTCGRKVSLPNKALELKDBETLSGKLEESPDGKLLKPTCGRKVS 507
Db 290 DKLINGKLEESPNKDGKLLKATCGMKVSIPTKALELMDMQTFKAEPPKPSAFEPATEMOKS 349
```

```
QY 508 LPNKALELKDBETLSGKLEESPDNDGKLLKPTCGRKVSLPNKALELKDBETLSGKLEESPDNDGKLLKPTCGRKVS 567
Db 350 VPKALELKNEQTLRAQ-----EILPSSSKQ 375
```

```
QY 568 KQDEENSWDFESFLETLQNDVCLPKATHQKEFTLSGKLEESPDGKLLKPTCGMKISL 627
Db 376 KYDEESSWDSSELCTVSKQDVCLPKAXHQKEIDKINGKLEESPVKDGKLLKATCGMKVSI 435
```

```
QY 628 PNKALELKDBETLSGKLEESPVKDGKLLKPTCGRKVSLPNKALELKDBETLSGKLEESPDNDGKLLKPTCGRKVS 687
Db 436 PTKALELMDMQTFKAEPPK-----PSAF-----EPAIEMQKSV-----PNKALELK 477
```

```
QY 688 QERDGIIRAPQDQTNKMPSELGRKEDTKSTSDSEIISVSDTQNYECLPEATYQKEIK 747
Db 478 NEQTLRADEILP-----SESKQDYESSWDSSELCTVSKQDVCLPKAXHQKEID 528
```

```
QY 748 TTNGKIBES-----PEKPSHFEPEATEMNSV 773
Db 529 KINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPPKPSAFEPATEMOKSV 588
```

```
QY 774 PNKGLEWKNQOTLRA----- 788
Db 589 PNKALELKNEQTLRAQMPFSESQKQKVEENSWDSSELRETVSQKDVCPKATHQKEMDK 648
```

```
QY 789 -----DSTTLKSLDALPSCERGERELKDNCEQITAKMEQMKNFVLOKELSEAKIK 842
Db 649 ISGKLESTLSKILDTVHSCERARELOKHCEQORTGMQKFKCVLKKLSEAKIK 708
```

```
QY 843 SQLENQKAKWEQELCSVRLPLNQBEERNNVDILKEKI 880
Db 709 SQLENQKAKWEQELCSVRLPLNQBEERNNVDILKEKI 735
```


487	Qy	FXAESPDKDGLLKPTCGRKVSLPNKALELKDRETLKAESPNDGLLKPTCGRKVSLPNKA	546
236	Db	FXAEPFGKPSAFEPATENQKSPVNPNALELKNEQTLRAD	273
547	Qy	LELKORETFKAAQMPFSSKOKDDENSWDPESFLETLQNQDVLCPKATHQEFDTLSGK	606
274	Db	LELKORETFKAAQMPFSSKOKDDENSWDPESFLETLQNQDVLCPKATHQEFDTLSGK	321
607	Qy	LBESPDKQGLLKPTCGMKISLPNKALELKDRETFKAEDSVSVSTFSLFGKPTTENSQST	666
322	Db	LEGSVPKQGLLKANGCMKVSIPTRKALELMDQTFKAEPPEK-PSAF	376
667	Qy	KVEEDFNLTKEGATKTVTGOERDIGHIERAPOQTNKMPSELGRKEDTKSTSEIL	726
377	Db	KVEEDFNLTKEGATKTVTGOERDIGHIERAPOQTNKMPSELGRKEDTKSTSEIL	414
727	Qy	SVSDTONVCELPATYQKEIKTTNGKIBES	756
415	Db	CETVSQKVCLPKAXHQEIDKINGKLEESPNDGFLKPCRMKVSIPTKALELMDQTF	474
757	Qy	PEKPSHFPATEMQNSVNPNGLEWKNQOTLRA	788
475	Db	KAEPPEKPSAFEPATENQKSPVNPNALELKNEQTLRADQMPFSSKQKXVEENSWDSESLR	534
789	Qy	PEKPSHFPATEMQNSVNPNGLEWKNQOTLRA	821
535	Db	ETVSQKDVCPKATHQKQEMDKISGKLEDDSTLSKILDTVHSCERARELQKHCEQTGM	594
822	Qy	EOMKNKFCVLQKELSEAKEIKSOLENQAKWEQELCSVRPLNQOEEKRNVDILKEI	880
595	Db	EOMKNKFCVLQKELSEAKEIKSOLENQAKWEQELCSVRPLNQOEEKRNVDILKEI	642

```

RESULT 11
US-09-433-826B-469
; Sequence 469, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433.826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-433-826B-469

Query Match      33.8%; Score 1751; DB 4; Length 650;
Best Local Similarity 49.6%; Pred. No. 1.6e-113;
Matches 386; Conservative 56; Mismatches 115; Indels 222; Gaps 9;

QY    187   LRPKETSEKFSWPAKERSRKITWEKEKTSVKTECVAGVTPNKTVLEKGTSNMIACTPK 246
       : |||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||::
DB     1    MSPAKETSEKTFWAAGPRKIANEKKETPVKTCGVARVTSNKTVKVLEKGSKWIACTPK 60

OY    247   ETSTKASTNDVDSSVEPFIISFLGTRTTIENSQCTVEEDFNLIATKI1SKSAAQNYTCLPDA 306

```

RESULT 13
US-09-834-759-469
; Sequence 469, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.

```
Db 475 KAEPPEKPSAFPALEMQKSPVNPKALELNQETLRADQMPPSESKQKXVEENSWDSLSR 534
Qy 789 -----DSTLSKILDALPSCERGRKELKONCEQITAKM 821
Db 535 ETVSQKDVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKM 594
Qy 822 EQMKNKFCVLQKELSEAKEIKSQLENQAKWQELCSVRLPLNQEBEERKRVNDILKEKI 880
Db 595 EQMKKKFCVLKKJSEAKEIKSQLENQKVKWQELCSVRF-----LTLMKMKI 642

RESULT 14
US-09-590-751A-469
; Sequence 469, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-590-751A-469

Query Match 33.8%; Score 1751; DB 4; Length 650;
Best Local Similarity 49.6%; Pred. No. 1.6e-113;
Matches 386; Conservative 56; Mismatches 115; Indels 222; Gaps 9;

Qy 187 LRPTKTSKFWPAKERSRKITWEEKTSVKTECVAGVTNKTVELEKGTSGNMIACTK 246
Db 1 MSPAKETSKEFTWAAGRPKIAWEKETPVKTGCVARVTSNKTVELEKGRSKMIACTK 60
Qy 247 ETSTKASTNVDSVSSVEPIPSLFGTRTIENSQCTKVEEDFNLAATKIISKSAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69
Qy 307 TYQDKITINHKLIEDQMPPSESKREDEEYSWDSGLSPFESSAKTQVCIPESMYQKWEIN 366
Db 70 -----DQRPPESEKQEBEYSCDSRSLFESSAKIQVCIPESYQKWEIN 115
Qy 367 REVEELPEKPSAFKPAVENQKTPVNPKALELNQETLRADQMPPSESKQKXVEENSWDS 426
Db 116 REVEELPEKPSAFKPAVENQKTPVNPKALELNQETLRADQMPPSESKQKXVEENSWDS 175
Qy 427 PCETVSQKDVYLPKATHQKEFTLQKLEESPVKDGLLKPTCGRKVSLNPKALELNQET 486
Db 176 LCETVSQKDVCLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDQTF 235
Qy 487 FXAESPDGGLLKPTCGRKVSLNPKALELNQETLRADQMPPSESKQKXVEENSWDSLSR 546
Db 236 FXAEPPEKPSAFEPATEMOMKSPVNPKALELNQETLRAD----- 273
Qy 547 LELKDRETFKAQMPPSESKQKDEENSWDFSEFTLLQNDVCLPKATHQKEFTLQK 606
Db 274 -----EILPSESKQKDEENSWDSSELCETVSQKDVCLPKATHQKEIDKINGK 321
Qy 607 LEESPDGGLLKPTCGMKVSLNPKALELNQETLRADQMPPSESKQKXVEENSWDSLSR 666
Db 116 REVEELPEKPSAFKPAVENQKTPVNPKALELNQETLRADQMPPSESKQKXVEENSWDS 175
```

```
Db 322 LEGSPVKDGLLKANCQMKVSIPTKALELMDQTFKAEPPEK-PSAF-----EPAIEMQKSV 376
Qy 667 KVEEDFNLTJTEGATKVTGQQRDGTIIRAPQDQTNKMPITSELGRKEDTKTSDSEII 726
Db 377 -----PNKALELNQETLRADLEILP-----SESKQKDEEENSWDSLSR 414
Qy 727 SVSDTQNYECLPEATYQKEIKTTNGKIEES----- 756
Db 415 CETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTF 474
Qy 757 -----PEKPSHPEPATEMOMKSPVNPKNGLKWNKQTLRA----- 788
Db 475 KAEPPEKPSAFEPATEMOMKSPVNPKALELNQETLRADQMPPSESKQKXVEENSWDSLSR 534
Qy 789 -----DSTLSKILDALPSCERGRKELKONCEQITAKM 821
Db 535 ETVSQKDVCPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKM 594
Qy 822 EQMKNKFCVLQKELSEAKEIKSQLENQAKWQELCSVRLPLNQEBEERKRVNDILKEKI 880
Db 595 EQMKKKFCVLKKJSEAKEIKSQLENQKVKWQELCSVRF-----LTLMKMKI 642

RESULT 15
US-09-551-621-469
; Sequence 469, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-469

Query Match 33.8%; Score 1751; DB 4; Length 650;
Best Local Similarity 49.6%; Pred. No. 1.6e-113;
Matches 386; Conservative 56; Mismatches 115; Indels 222; Gaps 9;

Qy 187 LRPTKTSKFWPAKERSRKITWEEKTSVKTECVAGVTNKTVELEKGTSGNMIACTK 246
Db 1 MSPAKETSKEFTWAAGRPKIAWEKETPVKTGCVARVTSNKTVELEKGRSKMIACTK 60
Qy 247 ETSTKASTNVDSVSSVEPIPSLFGTRTIENSQCTKVEEDFNLAATKIISKSAQNYTCLPDA 306
Db 61 ESSTKASAN----- 69
Qy 307 TYQDKITINHKLIEDQMPPSESKREDEEYSWDSGLSPFESSAKTQVCIPESMYQKWEIN 366
Db 70 -----DQRPPESEKQEBEYSCDSRSLFESSAKIQVCIPESYQKWEIN 115
Qy 367 REVEELPEKPSAFKPAVENQKTPVNPKALELNQETLRADQMPPSESKQKXVEENSWDS 426
Db 116 REVEELPEKPSAFKPAVENQKTPVNPKALELNQETLRADQMPPSESKQKXVEENSWDS 175
```


Qy 427 PCETVSQKDVLPKATHQKEPDTLSGKLESPVKGDLKPTCGRKVSIPNKALELKORET 486
 Db 176 LCETVSQKDVLPKATHQKEIDKINGKLEESPNKGLLKATCGMKVSIPTKALELKDQMT 235
 Qy 487 FKAESPDKDGLLKPTCGRKVSIPNKALELKORETILKAESPNDGDLKPTCGRKVSIPNKA 546
 Db 236 FKAESPDKDGLLKPTCGRKVSIPNKALELKORETILKAESPNDGDLKPTCGRKVSIPNKA 273
 Qy 547 LELKDRETFKAAQFPSPESKQKDEENSWDFESFLETLQNDVCLPKATHQKEPDTLSGK 606
 Db 274 -----EILPSESKQKDYEESSWDSSELCETVSQKDVCLPKATHQKEIDKINGK 321
 Qy 607 LEESPDKDGLLKPTCGMKISIPNKALELKORETILKAEDVSVSESTFSLFGKPTTENSQST 666
 Db 322 LEGSPVKGDLKCANCGMKVSIPTKALELMDQMTFKAEPPEK-PSAF-----EPAIEMQKSV 376
 Qy 667 KVEDFNLTTEGATKTVTGOERDGIIEAPQODTNKMTSELGRKEDTKSTDSEII 726
 Db 377 -----PNKALELKNEQTLRADEILP-----SESKQKDYEESSWDSSESL 414
 Qy 727 SVSDTONYECLPEATYQKEIKTTNGKIBES----- 756
 Db 415 CETVSQKDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMTF 474
 Qy 757 ----PEKPSHFEPATEMNSVPNKGLEWKNKQTLRA----- 788
 Db 475 KAEPPEKPSAFEPALIMQKSVNKALELKNEQTLRADQMPFSESKQKXVEENSWDSSESLR 534
 Qy 789 -----DSTLSKILDALPSCERGRELKKNCCQITAKM 821
 Db 535 ETVSQKDVCPKATHQKEMDKISGKLEDSTLSKILDVHSCERARELQKDHCEQRTCKM 594
 Qy 822 EQMKNKFCVLOKELSEAKEIKSOLENOKAKWEQELCSVRLPLNQEEKKRRNVILKEKI 880
 Db 595 EQMKNKFCVLOKELSEAKEIKSOLENOKAKWEQELCSVRLPLNQEEKKRRNVILKEKI 642

Search completed: February 19, 2005, 02:13:36
 Job time : 33.5342 secs

the eye blank (uspto)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	5173	100.0	1011	4	AAB84703	AAB84703 Amino aci
2	5173	100.0	1011	7	ADJ69751	ADJ69751 Human hea
3	5173	100.0	1239	6	ABJ37789	ABJ37789 Human tum
4	5173	100.0	1239	7	ADL93239	ADL93239 Human bre
5	5156	99.7	1225	6	ABR47547	ABr47547 Breast ca
6	3075	59.4	1341	4	AAB84702	AAB84702 Amino aci
7	3075	59.4	1341	5	ABJ05537	ABJ05537 Breast ca
8	3075	59.4	1341	6	ABU37784	ABU37784 Human tum
9	3075	59.4	1341	6	ABR47548	ABr47548 Breast ca
10	3075	59.4	1341	7	ADL93227	ADL93227 Human bre
11	3069	59.3	1349	6	ABJ37788	ABJ37788 Human tum
12	3069	59.3	1349	7	ADL93235	ADL93235 Human bre
13	2374	45.9	1013	6	ABU37783	ABU37783 Human tum
14	2374	45.9	1013	7	ADL93215	ADL93215 Human bre
15	2372.5	45.9	1095	4	AUU33357	Auu33357 Human bre
16	2372.5	45.9	1095	5	ABG78924	ABg78924 Human bre
17	2372.5	45.9	1095	6	ABJ37747	ABj37747 Human tum
18	2372.5	45.9	1095	7	ADL93315	ADL93315 Human bre
19	2367	45.8	1002	4	AUU33351	Auu33351 Human bre
20	2367	45.8	1002	5	ABG78918	ABg78918 Human bre
21	2367	45.8	1002	6	ABJ37741	ABj37741 Human tum
22	2367	45.8	1002	7	ADL93137	ADL93137 Human bre
23	2367	45.8	1002	8	ADE44427	Ade44427 Human bre
24	1758	34.0	561	6	ABJ37782	ABj37782 Human tum
25	1758	34.0	561	7	ADL93214	ADL93214 Human bre

Query Match		100.0%;	Score 5173;	DB 4;	Length 1011;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 1011; Conservative		0;	Mismatches	0;	Gaps 0;
QY	1	MVATLLSYGAVIEVONKASLTPLLLAIQKRSKOTVEFLTTKNNANAFNESCCTALMLAI	60		
DB	1	MVATLLSYGAVIEVONKASLTPLLLAIQKRSKOTVEFLTTKNNANAFNESCCTALMLAI	60		
QY	61	CEGSSIEVGMLLQQNVDFVPAEDIHGITAERYAAARGVNYIHQOLLEHIRKLPKPNQNTNP	120		
DB	61	CEGSSIEVGMLLQQNVDFVPAEDIHGITAERYAAARGVNYIHQOLLEHIRKLPKPNQNTNP	120		
QY	121	EGTSTGTPDEAAPLAERTPDTAESLLEKTPDPAARLVSGTSAKIOCLGAKSGKEPQSTE	180		
DB	121	EGTSTGTPDEAAPLAERTPDTAESLLEKTPDPAARLVSGTSAKIOCLGAKSGKEPQSTE	180		
QY	181	ETPRKILRTPKTESKFSHPAKERSRKITWEEKETSVKTECVAGVTPNKTEVLEKGTNNM	240		
DB	181	ETPRKILRTPKTESKFSHPAKERSRKITWEEKETSVKTECVAGVTPNKTEVLEKGTNNM	240		
QY	241	IACPTKETSTKASTNVDSVSPVIFSLFGTRTIENSQCTKVEEDFNLATKIISKSAANY	300		
DB	241	IACPTKETSTKASTNVDSVSPVIFSLFGTRTIENSQCTKVEEDFNLATKIISKSAANY	300		
QY	301	TCLPDATYQKDINTINHIEDQMPSEKREDEEYSDWSGSLFPSSAKTQVCIPESMYQ	360		
DB	301	TCLPDATYQKDINTINHIEDQMPSEKREDEEYSDWSGSLFPSSAKTQVCIPESMYQ	360		
QY	361	KVMEINREVEELPEKPSAKPAVEMOKTPNKAFLKNEQTLRAAQMPFSESKQKDDEN	420		
DB	361	KVMEINREVEELPEKPSAKPAVEMOKTPNKAFLKNEQTLRAAQMPFSESKQKDDEN	420		
QY	421	SWDSESCETVSQKDVYLPKATHQKEFDTLGSKLEESPVKDGLLKPTCGRKVSLPNKALE	480		
DB	421	SWDSESCETVSQKDVYLPKATHQKEFDTLGSKLEESPVKDGLLKPTCGRKVSLPNKALE	480		
QY	481	LKORETPKASPDKGLLKPTCGRKVSLPNKALELKDRETLKAESEPDNDGLLKPTCGRKV	540		
DB	481	LKORETPKASPDKGLLKPTCGRKVSLPNKALELKDRETLKAESEPDNDGLLKPTCGRKV	540		
QY	541	SLPNKALELKDRETFKAAQMPFSESKQKDDENSWDFESFLETLQNDVCLPKATHQKEF	600		
DB	541	SLPNKALELKDRETFKAAQMPFSESKQKDDENSWDFESFLETLQNDVCLPKATHQKEF	600		
QY	601	DTLSGKLESPDKGLLKPTCGMKISLPNKALELKDRETFKAEEDVSVSTFSLFGKPTT	660		
DB	601	DTLSGKLESPDKGLLKPTCGMKISLPNKALELKDRETFKAEEDVSVSTFSLFGKPTT	660		
QY	661	ENSOSTKVEEDFNLTITKEGATKTVTGOERDIGHIIRAPQDQTNKMPSTSELGRKEDTKST	720		
DB	661	ENSOSTKVEEDFNLTITKEGATKTVTGOERDIGHIIRAPQDQTNKMPSTSELGRKEDTKST	720		
QY	721	SDSEIISVSDTONYECLEPEATYQKEIKITNGKIEESPEKPSHPFEPATEMNSVFNKLEW	780		
DB	721	SDSEIISVSDTONYECLEPEATYQKEIKITNGKIEESPEKPSHPFEPATEMNSVFNKLEW	780		
QY	781	KNKQTLRADSTTLTKILDALPSCBGRRELKDNCEQITAKMEQMKNFVLOKELSAAKE	840		
DB	781	KNKQTLRADSTTLTKILDALPSCBGRRELKDNCEQITAKMEQMKNFVLOKELSAAKE	840		
QY	841	IKSQLENOKAKWEQELCSVRPLNQOEKRRNVILKEKIPERPEQLRKLKLEVKHQLQTL	900		
DB	841	IKSQLENOKAKWEQELCSVRPLNQOEKRRNVILKEKIPERPEQLRKLKLEVKHQLQTL	900		
QY	901	RIQDIELKSVTSNLSNQVSHTHSENDLPHENCMLKKEIAMLKLEVATLKHOHQVKNKYF	960		
DB	901	RIQDIELKSVTSNLSNQVSHTHSENDLPHENCMLKKEIAMLKLEVATLKHOHQVKNKYF	960		
QY	961	EDIKILQEKNAELQMTLKLKQKTVTKRASQVREQLKVLTAENTMTLSKLKE	1011		
DB	961	EDIKILQEKNAELQMTLKLKQKTVTKRASQVREQLKVLTAENTMTLSKLKE	1011		

RESULT 2

ADJ69751	ADU69751 standard; protein; 1011 AA.
XX	ADJ69751;
XX	AC
XX	AC
DT	06-MAY-2004 (first entry)
XX	Human heat mitochondrial protein as a therapeutic target SeqID1557.
DE	mitochondrial; human; screening assay; diabetes mellitus;
XX	Huntington's disease; osteoarthritis;
KW	Leber's hereditary optic neuropathy; LHON;
KW	mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW	myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW	neuroprotective; nontropic; antidiabetic; anticonvulsant; antiarthritic;
KW	osteopathic; ophthalmological; cytostatic.
XX	Homo sapiens.
OS	WO2003087768-A2.
XX	23-OCT-2003.
XX	04-APR-2003; 2003WO-US010870.
PF	12-APR-2002; 2002US-0372843P.
XX	17-JUN-2002; 2002US-038987P.
PR	20-SEP-2002; 2002US-0412418P.
XX	(MITO-) MITOKOR.
PA	(BUCK-) BUCK INST AGE RES.
XX	Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
PI	Warnock DE;
XX	WPI; 2003-845369/78.
DR	Identifying a mitochondrial target for drug screening assays and for
XX	treating diseases associated with altered mitochondrial function,
PT	comprises detecting a modified polypeptide in a sample and correlating
PT	with the disease.
PT	Claim 1; SEQ ID NO 1557; 180pp; English.
XX	This invention relates to novel mitochondrial targets that can be used
CC	for therapeutic intervention in treating a disease associated with
CC	altered mitochondrial function. Specifically, it refers to a method for
CC	identifying proteins of the human heart mitochondrial proteome that are
CC	useful for drug screening assays, as well as therapeutic targets. The
CC	present invention describes a method for identifying such proteins that
CC	can be used in the treatment of various diseases associated with altered
CC	mitochondrial function including diabetes mellitus, Huntington's disease,
CC	osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC	encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC	ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC	compositions have neuroprotective, nontropic, antidiabetic,
CC	anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC	cytostatic activities. This polypeptide sequence is a human heart
CC	mitochondrial protein of the invention.
XX	Sequence 1011 AA;
SQ	Query Match 100.0%; Score 5173; DB 7; Length 1011;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MVATLLSYGAVIEVONKASLTPLLLAIQKRSKOTVEFLTTKNNANAFNESCCTALMLAI 60
DB	1 MVATLLSYGAVIEVONKASLTPLLLAIQKRSKOTVEFLTTKNNANAFNESCCTALMLAI 60
QY	61 CEGSSIEVGMLLQQNVDFVPAEDIHGITAERYAAARGVNYIHQOLLEHIRKLPKPNQNTNP 120

Db 61 CEGSSEIVGMILLQONVDVFAEDIHGITAERYAAARGVNYIHQOLLEHIRKLPKNPONTNP 120
Qy 121 EGTSTGTPDEAAPLAERTPDPAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQST 180
Db 121 EGTSTGTPDEAAPLAERTPDPAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQST 180
Qy 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Db 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Qy 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Db 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Qy 301 TCLPDATYQKDIKTINHKEIDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQKDIKTINHKEIDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Qy 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTLRAAQMFSESKQKDEEN 420
Db 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTLRAAQMFSESKQKDEEN 420
Qy 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
Db 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480
Qy 481 LKDRFTFAESPDKGLLKPTCGRKVSLPNKALEKDRFTLKAESPNDGLLKPTCGRKV 540
Db 481 LKDRFTFAESPDKGLLKPTCGRKVSLPNKALEKDRFTLKAESPNDGLLKPTCGRKV 540
Qy 541 SLPNKALELKDRFTFAESPDKGLLKPTCGRKVSLPNKALEKDRFTLKAESPNDGLLKPTCGRKV 600
Db 541 SLPNKALELKDRFTFAESPDKGLLKPTCGRKVSLPNKALEKDRFTLKAESPNDGLLKPTCGRKV 600
Qy 601 DTLGKLEESPDKGLLKPTCGRKVSLPNKALELKDRFTFAESPDKGLLKPTCGRKV 660
Db 601 DTLGKLEESPDKGLLKPTCGRKVSLPNKALELKDRFTFAESPDKGLLKPTCGRKV 660
Qy 661 ENSQSTKVEEDFNLTITKEGATKTVTQGERDGIIERAPQDQTNKMPSELGKEDTKST 720
Db 661 ENSQSTKVEEDFNLTITKEGATKTVTQGERDGIIERAPQDQTNKMPSELGKEDTKST 720
Qy 721 SDEIISVSDTQNYECLPEATYQKEIKTTNGKIESEPEKPSHFEPATENQNSVPNKGLEW 780
Db 721 SDEIISVSDTQNYECLPEATYQKEIKTTNGKIESEPEKPSHFEPATENQNSVPNKGLEW 780
Qy 781 KXKQTLRADSTTLTKILDALPSCERGLKNDKNCQITAKMEQKNNKFCVLOKELSEAKE 840
Db 781 KXKQTLRADSTTLTKILDALPSCERGLKNDKNCQITAKMEQKNNKFCVLOKELSEAKE 840
Qy 841 IKSQLENQKAKWEQELCSVRLPLNQBEEKRRNVVDILKEKIRPEEQLRKKLEVKHQLEQTL 900
Db 841 IKSQLENQKAKWEQELCSVRLPLNQBEEKRRNVVDILKEKIRPEEQLRKKLEVKHQLEQTL 900
Qy 901 RIQDIELKSVTNLQNVSHSENDLPHENCMLKKEIAMLKLEVATLKHOHVKENKYF 960
Db 901 RIQDIELKSVTNLQNVSHSENDLPHENCMLKKEIAMLKLEVATLKHOHVKENKYF 960
Qy 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011
Db 961 EDIKILOEKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011

RESULT 3

ABJ37789
ID ABJ37789 standard; protein; 1239 AA.

XX AC ABJ37789;

XX DT 15-MAY-2003 (first entry)

XX DE Human tumour-related protein - SEQ ID No 577.

XX

KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
OS Homo sapiens.
PN WO200283956-A1.
XX 24-OCT-2002
XX 15-APR-2002; 2002WO-US012378.
XX 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedwick JS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
XX New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX Example 13; Page 359-363; 375pp; English.
XX The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 1239 AA;

Query Match 100.0%; Score 5173; DB 6; Length 1239;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSQTVFELLTKNANAFNESCCTALMLAI 60
Db 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSQTVFELLTKNANAFNESCCTALMLAI 60
Qy 61 CEGSSEIVGMILLQONVDVFAEDIHGITAERYAAARGVNYIHQOLLEHIRKLPKNPONTNP 120
Db 61 CEGSSEIVGMILLQONVDVFAEDIHGITAERYAAARGVNYIHQOLLEHIRKLPKNPONTNP 120
Qy 121 EGTSTGTPDEAAPLAERTPDPAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQST 180
Db 121 EGTSTGTPDEAAPLAERTPDPAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQST 180
Qy 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Db 181 ETPRKILRPTKETSEKFSWPAKERSRKITWEEKETSVKTECVAGVTPNKTVELEKTSNM 240
Qy 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Db 241 IACPKTETSKASTNVDSVSSVEPIFSLFGTRTITENSQCTKVBEEDFNLAATKIIKSAAQNY 300
Qy 301 TCLPDATYQKDIKTINHKEIDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQKDIKTINHKEIDQMPPSEKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
Qy 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTLRAAQMFSESKQKDEEN 420
Db 361 KWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTLRAAQMFSESKQKDEEN 420
Qy 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLSGKLEESPVKDGLLKPTCGRKVSLPNKALE 480

Db 421 SWDSSEPCETVSQKQVYLPKATHQKEFDLTKGLSESPVKDGLLKTGCRKVSLEPNKALE 480
QY 481 LKDRETFKAESPDKDGLLKTGCRKVSLEPNKALELKDRETLKAESPNDGLLKTGCRKV 540
Db 481 LKDRETFKAESPDKDGLLKTGCRKVSLEPNKALELKDRETLKAESPNDGLLKTGCRKV 540
QY 541 SLPNKALELKDRETFKAQMPFSESKQKDDDENSWDFSEFLTLQNDVCLPKATHQKEF 600
Db 541 SLPNKALELKDRETFKAQMPFSESKQKDDDENSWDFSEFLTLQNDVCLPKATHQKEF 600
QY 601 DTLSGKLESPDKDGLLKTGCMKISLPNKALELKDRETFKAEDVSSVESTSLFGKPTT 660
Db 601 DTLSGKLESPDKDGLLKTGCMKISLPNKALELKDRETFKAEDVSSVESTSLFGKPTT 660
QY 661 ENSQSTKVEEDFNLTTKGATKTVTGQQRDGIIGIERAPQDQTNKMPSELGRKEDTKST 720
Db 661 ENSQSTKVEEDFNLTTKGATKTVTGQQRDGIIGIERAPQDQTNKMPSELGRKEDTKST 720
QY 721 SDEIISVSQDQNYECLPEATYQKEIKTTNGKIEESPEKPSHFEPATEMONSPVNGLEW 780
Db 721 SDEIISVSQDQNYECLPEATYQKEIKTTNGKIEESPEKPSHFEPATEMONSPVNGLEW 780
QY 781 KKKQTLRADSTTLTKILDALPSCERGRELKONCEQITAKMEQMKNFVLOKELSEAKE 840
Db 781 KKKQTLRADSTTLTKILDALPSCERGRELKONCEQITAKMEQMKNFVLOKELSEAKE 840
QY 841 IKSQLEQKAKWEQELCSVRLPLNQBEKRRNVDIILKEKIRPBEQLRKKLVKHQLEQTL 900
Db 841 IKSQLEQKAKWEQELCSVRLPLNQBEKRRNVDIILKEKIRPBEQLRKKLVKHQLEQTL 900
QY 901 RIQDIELKSVTSNLQVSHSENDLFHENCMLKKEIAMLVATLKHQOVKENYF 960
Db 901 RIQDIELKSVTSNLQVSHSENDLFHENCMLKKEIAMLVATLKHQOVKENYF 960
QY 961 EDIKILQEKNAELQMTLKLKQKTVTKRASQYRQLKVLTAENTMLTSKLE 1011
Db 961 EDIKILQEKNAELQMTLKLKQKTVTKRASQYRQLKVLTAENTMLTSKLE 1011
RESULT 4
ADL93239
ID ADL93239 standard; protein; 1239 AA.
XX AC ADL93239;
XX DT 20-MAY-2004 (first entry)
XX DE Human breast cancer-associated polypeptide #54.
XX KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX OS Homo sapiens.
XX PN US2003166022-A1.
XX PD 04-SEP-2003.
XX PF 15-APR-2002; 2002US-00124805.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.

(CORI-) CORIXA CORP.
Houghton RL, Sleath PR, Persing DH;
WPI; 2003-874918/81.
N-PSDB; ADL93238.
An isolated oncogenic polypeptide useful for preventing, diagnosing and treating breast cancer.
Example 13; SEQ ID NO 577; 294pp; English.
The invention relates to an isolated breast cancer-associated polypeptide. The polypeptide may be used for the diagnosis and treatment of breast cancers. The methods are useful for detecting the presence of a cancer in a patient and treating a cancer in a patient. The present sequence represents the amino acid sequence of a human breast cancer-associated polypeptide.
Sequence 1239 AA;
Query Match 100.0%; Score 5173; DB 7; Length 1239;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSKQTVFELTKNANANAFNEKCTALMLAI 60
Db 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSKQTVFELTKNANANAFNEKCTALMLAI 60
QY 61 CEGSSEIVGMLLQNVVFAEDIHGITAERYAAGVNYIHQOLLEHIRKLPKNQNTNP 120
Db 61 CEGSSEIVGMLLQNVVFAEDIHGITAERYAAGVNYIHQOLLEHIRKLPKNQNTNP 120
QY 121 EGTSTGTDPDEAPLAERTPDTAESILEKTPDEAARLVEGTSAKIOCLGKATSGKEQSTE 180
Db 121 EGTSTGTDPDEAPLAERTPDTAESILEKTPDEAARLVEGTSAKIOCLGKATSGKEQSTE 180
QY 181 ETPRKILRPKTETSEKFSWPAKERSKTIWEEKTSVKTECVAGVTPNKTEVLEKGTSM 240
Db 181 ETPRKILRPKTETSEKFSWPAKERSKTIWEEKTSVKTECVAGVTPNKTEVLEKGTSM 240
QY 241 IACPTKETSTKASTNVDSVSSVEPIFSLFGTRIENSQCTKVEEDFNLATKIISKAAQNY 300
Db 241 IACPTKETSTKASTNVDSVSSVEPIFSLFGTRIENSQCTKVEEDFNLATKIISKAAQNY 300
QY 301 TCLPDATYQDKIKTINHIEDQMPFSESKREDEEYSWDSGLFESSAKTQVCIPESMYQ 360
Db 301 TCLPDATYQDKIKTINHIEDQMPFSESKREDEEYSWDSGLFESSAKTQVCIPESMYQ 360
QY 361 KYWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTLRAAQMPFSESKQKDDEN 420
Db 361 KYWEINREVEELPEKPSAFKPAVEMQKTPNKAFAELKNEQTLRAAQMPFSESKQKDDEN 420
QY 421 SWDSSEPCETVSQKQVYLPKATHQKEFDLTKGLSESPVKDGLLKTGCRKVSLEPNKALE 480
Db 421 SWDSSEPCETVSQKQVYLPKATHQKEFDLTKGLSESPVKDGLLKTGCRKVSLEPNKALE 480
QY 481 LKDRETFKAESPDKDGLLKTGCRKVSLEPNKALELKDRETLKAESPNDGLLKTGCRKV 540
Db 481 LKDRETFKAESPDKDGLLKTGCRKVSLEPNKALELKDRETLKAESPNDGLLKTGCRKV 540
QY 541 SLPNKALELKDRETFKAQMPFSESKQKDDDENSWDFSEFLTLQNDVCLPKATHQKEF 600
Db 541 SLPNKALELKDRETFKAQMPFSESKQKDDDENSWDFSEFLTLQNDVCLPKATHQKEF 600
QY 601 DTLSGKLESPDKDGLLKTGCMKISLPNKALELKDRETFKAEDVSSVESTSLFGKPTT 660
Db 601 DTLSGKLESPDKDGLLKTGCMKISLPNKALELKDRETFKAEDVSSVESTSLFGKPTT 660
QY 661 ENSQSTKVEEDFNLTTKGATKTVTGQQRDGIIGIERAPQDQTNKMPSELGRKEDTKST 720
Db 661 ENSQSTKVEEDFNLTTKGATKTVTGQQRDGIIGIERAPQDQTNKMPSELGRKEDTKST 720

QY 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHFEPATEMNSVFNKGLW 780
 Db 721 SDSEIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHFEPATEMNSVFNKGLW 780
 QY 781 KKKQTLRADSTLTKILDALPSCERGLKKDNCQITAKMEQKNKFCVLQKELSEAKE 840
 Db 781 KKKQTLRADSTLTKILDALPSCERGLKKDNCQITAKMEQKNKFCVLQKELSEAKE 840
 QY 841 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
 Db 841 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
 QY 901 RIQDIELKSVTSNLSNOVSHTHSENDLFHENCMLKKEIAMLKLEVATLKHQHVKNKYF 960
 Db 901 RIQDIELKSVTSNLSNOVSHTHSENDLFHENCMLKKEIAMLKLEVATLKHQHVKNKYF 960
 QY 961 EDIKILOKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011
 Db 961 EDIKILOKNAELQMTLKLKQKTVTKRASQYREQLKVLTAENTMLTSKLKE 1011

RESULT 5

ABR47547
 ID ABR47547 standard; protein; 1225 AA.

XX ABR47547;

DT 12-JUN-2003 (first entry)

DE Breast cancer associated protein sequence SEQ ID NO:332.

KW Human; breast cancer; cytostatic; gene therapy.

XX Homo sapiens.

XX WO2003004989-A2.

XX 16-JAN-2003.

XX 21-JUN-2002; 2002WO-US019669.

XX 21-JUN-2001; 2001US-0299887P.

PR 27-JUN-2001; 2001US-0301572P.

PR 18-JUL-2001; 2001US-0306501P.

PR 25-SEP-2001; 2001US-0325002P.

PR 05-MAR-2002; 2002US-0362585P.

PR 14-MAY-2002; 2002US-0380391P.

XX (MILL-) MILLENIUM PHARM INC.

XX Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;

PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;

PI Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX WPI; 2003-210381/20.

DR N-ESDB; ACC50245.

XX Breast cancer diagnosis or treatment by comparing the level of expression

PT of a marker in a patient sample with that in the control non-breast

PT cancer sample.

XX Claim 1; SEQ ID NO 332; 128pp; English.

PS The present invention describes a method for assessing whether a patient

CC is afflicted with breast cancer. The method comprises comparing the level

CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and

CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1225 AA;

Query Match 99.7%; Score 5156; DB 6; Length 1225;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1008; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVATLLSGAVIEQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESCCTALMLAI 60
 Db 154 MVATLLSGAVIEQNKASLTPLLLAIQKRSKQTVFELLTKNANANAFNESCCTALMLAI 213
 QY 61 CEGSSEIVGMLLQONVDVFAEDIHGITAERVAARGVNYIHQOLLEHRLKLPKNQNTNP 120
 Db 214 CEGSSEIVGMLLQONVDVFAEDIHGITAERVAARGVNYIHQOLLEHRLKLPKNQNTNP 273
 QY 121 EGTSTGTPDEAAPLAERTPDTAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQSTE 180
 Db 274 EGTSTGTPDEAAPLAERTPDTAESLLEKTPDEAARLVEGTSAKIQCLGKATSGKFEQSTE 333
 QY 181 ETPRKILRPYKTSKFSWPAKERSRKITWEKETSVKTECVAGVTPNKTVELEKGTSM 240
 Db 334 ETPRKILRPYKTSKFSWPAKERSRKITWEKETSVKTECVAGVTPNKTVELEKGTSM 393
 QY 241 IACPTKETSTKASTNVDSVVEPIPSLFGTITENSQCTKVEEDFNLATKIISKSAQNY 300
 Db 394 IACPTKETSTKASTNVDSVVEPIPSLFGTITENSQCTKVEEDFNLATKIISKSAQNY 453
 QY 301 TCLPDATYQKDIKTINHKKIEDQMPFSESKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 360
 Db 454 TCLPDATYQKDIKTINHKKIEDQMPFSESKREDEEYSWDSGSLFESSAKTQVCIPESMYQ 513
 QY 361 KWEINREVEELPEKPSAFKPAVEQKQTPNKAELKNEQTLRAAQMPFSESKQKDDEN 420
 Db 514 KWEINREVEELPEKPSAFKPAVEQKQTPNKAELKNEQTLRAAQMPFSESKQKDDEN 573
 QY 421 SWDSSEPCETVSQKDVYLPKATHQKEFDTLGKLEESPVKDGLLKPTCGRKVSLENKALE 480
 Db 574 SWDSSEPCETVSQKDVYLPKATHQKEFDTLGKLEESPVKDGLLKPTCGRKVSLENKALE 633
 QY 481 LKDRETFKAESPDKDGLLKPTCGRKVSLENKALEKDRETLKAESPNDGLLKPTCGRKV 540
 Db 634 LKDRETFKAESPDKDGLLKPTCGRKVSLENKALEKDRETLKAESPNDGLLKPTCGRKV 693
 QY 541 SLPNKALELKDRETFKAQAMPFSESKQKDDENSWDFESFLETLQNDVCLPKATHQKEF 600
 Db 694 SLPNKALELKDRETFKAQAMPFSESKQKDDENSWDFESFLETLQNDVCLPKATHQKEF 753
 QY 601 DTLGKLEESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESTFSFGKPTT 660
 Db 754 DTLGKLEESPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESTFSFGKPTT 813
 QY 661 ENSQSTKVEEDFNLTITKEGATKTVTGQERDIGIIERAPODQTNKMPSTSELGRKEDTKST 720
 Db 814 ENSQSTKVEEDFNLTITKEGATKTVTGQERDIGIIERAPODQTNKMPSTSELGRKEDTKST 873
 QY 721 SDSIIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHFEPATEMNSVFNKGLW 780
 Db 874 SDSIIISVSDTQNYECLPEATYQKEIKTTNGKIBESPEKPSHFEPATEMNSVFNKGLW 933
 QY 781 KKKQTLRADSTLTKILDALPSCERGLKKDNCQITAKMEQKNKFCVLQKELSEAKE 840
 Db 934 KKKQTLRADSTLTKILDALPSCERGLKKDNCQITAKMEQKNKFCVLQKELSEAKE 993
 QY 841 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 900
 Db 994 KKSQLENQAKWEQELCSVRLPLNQEEKKRRNVDILKEKIRPEEQLRKKLEVKHOLEQTL 1053
 QY 901 RIQDIELKSVTSNLSNOVSHTHSENDLFHENCMLKKEIAMLKLEVATLKHQHVKNKYF 960
 Db 1054 RIQDIELKSVTSNLSNOVSHTHSENDLFHENCMLKKEIAMLKLEVATLKHQHVKNKYF 1113

Tue Feb 22 09:49:15 2005

QY	961	EDIKILOEKNAELQMTLKLKQKVTTRKASQYREQLKVLTAENTMLTSLKE	1011
Db	1114	EDIKILOEKNAELQMTLKLKQKVTTRKASQYREQLKVLTAENTMLTSLKE	1164
RESULT 6			
AA	84702	AAB84702 standard; protein; 1341 AA.	
XX	AC	AAB84702;	
XX	DT	17-SEP-2001 (first entry)	
XX	DE	Amino acid sequence of a human cancer associated antigen.	
XX	KW	Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.	
XX	OS	Homo sapiens.	
XX	PN	WO200147959-A2.	
XX	PD	05-JUL-2001.	
XX	PF	29-NOV-2000; 2000WO-US042334.	
XX	PR	30-NOV-1999; 99US-00451739.	
XX	PR	24-OCT-2000; 2000US-00602362.	
XX	PA	(LUDW-) LUDWIG INST CANCER RES.	
XX	PA	(SLOK) SLOAN KETTERING INST CANCER RES.	
XX	PA	(CORR) CORNELL RES FOUND INC.	
XX	PI	Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;	
XX	DR	WPI; 2001-441706/47.	
XX	PT	Isolated cancer associated nucleic acid molecule identified by SEREX	
XX	PT	(serological identification of antigens by recombinant expression	
XX	PT	cloning) technique, useful in nucleic acid based therapies to treat	
XX	PT	cancer.	
XX	PS	Claim 83; Page 53-57; 62pp; English.	
XX	CC	The present sequence represents a human cancer associated antigen. The	
XX	CC	sequence was identified using probes derived from the INGI gene. The INGI	
XX	CC	gene is a tumour suppressor candidate gene. The cancer associated antigen	
XX	CC	polynucleotides and polypeptides are useful for screening for the	
XX	CC	possible presence of a pathological condition in a subject such as	
XX	CC	cancer. The cancer associated antigen polypeptides are useful for	
XX	CC	producing vaccines	
XX	SQ	Sequence 1341 AA;	
Query Match			
Best Local Similarity 59.4%; Score 3075; DB 4; Length 1341;			
Matches 664; Conservativity 57.4%; Pred. No. 3.8e-199;			
Indels 272; Gaps 12;			
QY	1	MYATLLSGAVIEVQNKASLPLLLAIOKRSKQTVFLLTKNANAFNKSCTALMLAI	60
Db	98	VVAKLLSHGAVIEVHNKASLPLLLSITRSEQIVFLLIKNANAVNKYKCTALMLAV	157
QY	61	CEGSSEIVGMLLQONVDVFAEDIHGITAERYAARGVNYIHQOLLEHRLKLPKQNTNP	120
Db	158	CHGSSEIVGMLLQONVDVFAADICGVTAHYAVTCGFHHIHEQIMEYIRKLSKQNTNP	217
QY	121	EGTSTGTDPDEAPLA-----BRTPTAESLLEKTPDEAARLVE	158
Db	218	EGTSAGTDPDEAPLAERTPTDAESLVEKTPDEAALVETPTDAESLVEKTPDEAASLVE	277
QY	159	GTSAKIOCLGKATSGFQESTETPRKILRPTKETSEKFSWP-----	200
Db	278	GTSDKIQCLEKATSGFQESAETPREITSPAKETSEKFTWPAKGRPKRIAWKEDTTPR	337

RESULT 7
 ABJ05537
 ID ABJ05537 standard; protein; 1341 AA.
 XX
 AC ABJ05537;
 XX
 DT 14-NOV-2002 (first entry)

XX Breast cancer-associated protein 2.
XX Breast cancer; breast cancer-associated gene sequence; drug development;
KW pharmacogenetics; biosensor development.
XX Unidentified.
OS
XX WO200259377-A2.
XX 01-AUG-2002.
XX 24-JAN-2002; 2002WO-US002242.
XX 24-JAN-2001; 2001US-0263965P.
XX 02-FEB-2001; 2001US-0265928P.
XX 09-APR-2001; 2001US-00829472.
XX 09-APR-2001; 2001US-0282698P.
XX 04-MAY-2001; 2001US-0288590P.
XX 29-MAY-2001; 2001US-0294443P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Mack DH, Gish KC, Afar D;
XX WPI; 2002-583738/62.
XX N-PSDB; ABT07694.
XX
XX Detecting a breast cancer-associated transcript in a patient's cell,
XX useful for diagnosing breast cancer, comprises contacting a biological
XX sample with a polynucleotide that selectively hybridizes with breast
XX cancer nucleic acids.
XX
XX Disclosure; Page 348-349; 414pp; English.
XX
XX The invention comprises a method of detecting a breast cancer-associated
XX transcript in a cell from a patient. The method of the invention involves
XX contacting a biological sample from the patient with a nucleotide that
XX hybridizes to one of the 69 breast cancer-associated gene sequences shown
XX in the specification. The method of the invention is useful in the
XX diagnosis or prognosis of breast cancer, and for detecting genes that are
XX up or down-regulated in breast cancer cells. Genes identified by the
XX method of the invention can be used in diagnostic purposes and also as
XX targets for screening for therapeutic compounds that modulate breast
XX cancer (e.g. hormones or antibodies). Identification of genes that are
XX over or under expressed in breast cancer can additionally provide high-
XX resolution, high-sensitivity datasets which can be used in the areas of
XX diagnostics, therapeutics, drug development, pharmacogenetics, protein
XX structure and biosensor development. Amino acid sequences ABJ05536 -
XX associated genes of the invention
XX
XX Sequence 1341 AA;
XX
XX Query Match 59.4%; Score 3075; DB 5; Length 1341;
XX Best Local Similarity 57.4%; Pred. No. 3.8e-199;
XX Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;
XX
XX 1 MVATLLSGAVTEVONKASLTLLLLAIOKRSQOTVEFLTKVANANAFNESCXTALMLAI 60
XX 98 VVAKLLSHGAVTEVHNKASLTLLLSITKRSQIIVEFLLIKANANANVNYKCTALMLAV 157
XX
XX 61 CEGSSEIVGMLLQNVVDVFAEDIHGTAEVYAAAGVNYIHOOLLEHRIKLPKNQNTNP 120
XX 158 CHGSSEIVGMLLQNVVDVFAADICGVTAEHYAVTCGFHHIEQIWEYIRKLSKNQNTNP 217
XX
XX 121 EGTSTGTDPDEAAPLA-----ERTPDATSLLEKTPDEAARLVE 158
XX 218 EGTSGTDPDEAAPLAERTPDATSLVEKTPDEAAPLVERTPDATSLVEKTPDEAASLVE 277
XX
XX 159 GTSAKIOCLGKATSGKFGQSTETETPRKILRPHKETSEKFSWP----- 200
XX 278 GTSKIOCLEKATSGKFGQSAETPREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPR 337

QY 201 -----AKERSRKITWEEKETSVKTECVAGVTPNKTVELEKGTSMIACP 244
DB 338 EIMSPAKETSEKFTWAAKGRPKIAWEKSETPVKTCVAVRTSNKTKVLEKGRSKMIACP 397
QY 245 TKETSTKASTNVDSVSRPIFSLFGRTRTIENSQCTKVBEEDFNLAITKIISKSAQAQNTCLP 304
DB 398 TKESSTKASAN----- 408
QY 305 DATYQDKDITINHKTEDQMFPSKREDEEYSWDSGLFESSAKTOVCIPESMYOKVME 364
DB 409 -----DQRFPSKQDEEYSCDSRSLFESSAKIQVCIPSIYOKVME 452
QY 365 INREVEELPEKPSAFKPAVEMQKTPVNFAPKELKNEOTLRAAQMFPSKQKDEENSWDS 424
DB 453 INREVEEPPKPSAFKPALEMQNSVPNFAPKELKNEOTLRAQMFPSKQKDEENSWDS 512
QY 425 ESPCTVSQKDVYLPKATHQKEFDTLGKLESPVKDGLLKPTCKGRKVSLPNKALELKOR 484
DB 513 ESLCETVSQKDVCLPKATHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELKD 572
QY 485 ETKFAESPDKGLLKPTCKGRKVSLPNKALELKADRETLEKESPDNDGLLKPTCKGRKVSLEN 544
DB 573 QTFKAPPPKPSAFEP-----ATEMOK-----SVEN 598
QY 545 KALELKADRETLEKESKQKDEENSWDSFLETLQNDVCLPKATHQKEFDTLG 604
DB 599 KALELKNEOTWRADEILFESKQKDEENSWDSILCETVSQKDVCLPKAAHQKEIDKIN 658
QY 605 GKLESPKDGKLLKPTCGMKISLPNKALELKADRETLEKESVSTFSLFKQKPTTENSQ 664
DB 659 GKLESPVKDGLLKANKCGMKVSIPTKALELMDQTFKAEPPK-PSAF-----EPATEMOK 713
QY 665 STKVEEDFNLTTEKATVTTCQOERDICIETARPDQTNKMTSELGKEDTKSTDSSE 724
DB 714 SV-----PNKALELKNEOTLRADEILP-----SESKQKDEENSWDS 751
QY 725 IISVSDTQNYECLPEATYQKEIKTTNGKIEES----- 756
DB 752 SILCETVSQKDVCLPKATHQKEIDKINGKLESPNDNGFLKAPCRMKVSIPTKALELMDMQ 811
QY 757 -----PKSHFEPATEMQNSVPNKGKWKQTLRA----- 788
DB 812 TFKAPPEKPSAFEPKPSVNPVNPKALELKNEOTLRAQMFPSKQKDEENSWDS 871
QY 789 -----DSTLTKILDALPSCERGBELKXKDCOEITA 819
DB 872 LRETVSQKDVCPKATHQKEMDKISGKLESDSTLSKILDTVHSCBRARELQKHCEQRTG 931
QY 820 KMEQMKKFCVQLQKELSEAKIISQLENOKAKWEQELCSVRLPLNQEBEKKRNVDILKEK 879
DB 932 KMEQMKKFCVQLKXKSEAKIISQLENOKAKWEQELCSVRLPLNQEBEKKRNVDILNEK 991
QY 880 IRP-----EQLRKKLEVKHQLEOTLRIQDIELKSVTSLNQVSHTHSENDLFHENCML 934
DB 992 IREELGRITEQHRKELEVKQLEQALRIQDIELKSVESNLNQVSHTHENENYLLHENCML 1051
QY 935 KKETAMKLEVALTKHOKVKNKYFEDIKLOENKNAELOMTLKLKOKTVTKRASQYREQ 994
DB 1052 KKETAMKLEVALTKHOKVKNKYFEDIKILKEKNAELOMTLKLKESLTKRASQYSQ 1111
QY 995 LKVLTAENTMLTSLKKE 1011
DB 1112 LKVLTAENTMLTSLKKE 1128
RESULT 8
ABJ37784
ID ABJ37784 standard; protein; 1341 AA.
XX
AC ABJ37784;
XX
DT 15-MAY-2003 (first entry)

XX Human tumour-related protein - SEQ ID NO 565.
DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
XX Human; breast cancer; cancer; immune response stimulation.
KW tumour; breast cancer; cancer; immune response stimulation.
XX Homo sapiens.
OS
XX WO200283956-A1.
XX 24-OCT-2002.
XX 15-APR-2002; 2002WO-US012378.
XX 13-APR-2001; 2001US-00834759.
XX 07-DEC-2001; 2001US-00007805.
XX 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Panger GR;
PI Vedvick TS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
XX New polypeptide and polynucleotide useful for stimulating and/or
XX expanding T cells specific for a tumor protein and treating breast
XX cancer.
XX Example 9; Page 342-346; 375pp; English.
XX The invention comprises a method of stimulating and/or expanding T cells
XX specific for a tumour protein. The invention further comprises human
XX nucleic acids and proteins that are associated with tumours (e.g. breast
XX cancer). The method and sequences of the invention are useful for
XX stimulating and/or expanding T cells specific for a tumour protein,
XX detecting the presence of cancer, stimulating an immune response in a
XX patient and treating breast cancer. The present amino acid sequence
XX represents a human tumour-related protein
XX
XX Sequence 1341 AA;
Query Match. 59.4%; Score 3075; DB 6; Length 1341;
Best Local Similarity 57.4%; Pred. No. 3.8e-199;
Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;
1 MVATLLSYGAVIEVQNKASLTPLLLATQKRSQTVFELLTKNANAFNESCCTALMLAI 60
98 VVAKLLSHGAVIEVHNKASLTPLLLSITKSEQIVFELLKNANAVNKYKCTALMLAV 157
61 CEGSEIVGMLLQNVDFVAFDIGHITAEVYAAAGVNYIHQQLLEHRIKLPGNPQNTNP 120
158 CHGSEIVGMLLQNVDFVAFDIGHITAEVYAVTCGFHHIHEQIMEYIRKLSKNHNTNP 217
121 EGTSTGTDPDEAPLA-----BRTDPTAESLLEKTPDEAARLVE 158
218 EGTSGTDPDEAPLAERTPTDAESLVEKTPDEAAPLVERTPDAESLVEKTPDEAARLVE 277
159 GTSAKIQCLGKATSKGFQSTQTEETPKILRPTKETSEKFSWP----- 200
278 GTSKIQCLGKATSKGFQSTQTEETPKILRPTKETSEKFSWP----- 337
201 -----AKERSKITWEKETSVKTECVAGVTPNKTVELEKGTSMNIACP 244
338 EIMSPAKETSEKFTWAAKGRPKIAWEKKEKTPVKTGCVARVTSNKTVELEKGRSMIACP 397
245 TKETSTKASTNVDSVVEFIFSLFGTRIENSQCTKVEEDFNLA TKIISKSAQAQNTCLP 304
398 TKESSTKASAN----- 408
305 DATYQDKINTINH KIDQMPFSGESKREDEEYSWDSGLSFESSAKTQVCIPESMTQKWE 364

Db 409 -----DQRFPSKQDEEYSCDSRSLFESSAKIQVCIPESIVQKWE 452
Qy 365 INREVELPEKPSAFKPAVEMQKTPVNAKAFELNKSQTTLRAAQMPFSESQKQDEENSWDS 424
Db 453 INREVEEPKPSAFKPAIEMQNSVPNAKAFELNKSQTTLRAQMPFSESQKQDEENSWDS 512
Qy 425 EPCETVSQKDVVLPKATHQKEFDTLISGKLEESPVKDGLLKPTCGRKVSLPKALELKDOR 484
Db 513 ESLCETVSQKDVVLPKATHQKEFDTLISGKLEESPVKDGLLKPTCGRKVSLPKALELKDOR 572
Qy 485 ETPKAESPDKDGLLKPTCGRKVSLPKALELKDORRETLKAESPNDGLLKPTCGRKVSLPN 544
Db 573 QTFKARPPGKPSAFEP-----ATEMOK-----SVEN 598
Qy 545 KALELKDORRETFKAAQMPFSESQKQDEENSWDSFLETLQNDVCLPKATHQKEFDTLIS 604
Db 599 KALELKDORRETFKAAQMPFSESQKQDEENSWDSFLETLQNDVCLPKATHQKEFDTLIS 658
Qy 605 GKLEESPDKGLLKPTCGMKISLPNAKALELKDORRETFKAEVSVSVSTFSLFGPPTTENSQ 664
Db 659 GKLEESPDKGLLKPTCGMKISLPNAKALELKDORRETFKAEVSVSVSTFSLFGPPTTENSQ 713
Qy 665 STKVEEDFNLTTEGATKTVTGQOERDIGHIIRAPQDQTNKMTSELGRKEDTKSTSDSE 724
Db 714 SV-----PNKALELKDORRETFKAEVSVSVSTFSLFGPPTTENSQ 751
Qy 725 IISVSDTQNYECLPEATYQKEIKTTNGKIBES----- 756
Db 752 SLCTETVSQKDVVLPKATHQKEFDTLISGKLEESPVKDGLLKPTCGRKVSLPKALELMDQ 811
Qy 757 -----PEKPSHFEPAEMQNSVPNAKALELKDORRETFKAEVSVSVSTFSLFGPPTTENSQ 788
Db 812 TFKAEPEKPSAFKPAIEMQNSVPNAKALELKDORRETFKAEVSVSVSTFSLFGPPTTENSQ 871
Qy 789 -----DSTTLISKILDALPSCERGERELKKNCEQITA 819
Db 872 LRETVSQKDVVLPKATHQKEFDTLISGKLEESPVKDGLLKPTCGRKVSLPKALELMDQ 931
Qy 820 KMEQMKNFVQLQKELSEAKIKSLENQKAKVQELCSVRLPLNOBEKERNVDILKEK 879
Db 932 KMEQMKNFVQLQKELSEAKIKSLENQKAKVQELCSVRLPLNOBEKERNVDILKEK 991
Qy 880 IRP-----EEQLRKLEVKHOLEQTLRIODIELKSVTSLNNOVSHTHSENDFHENCML 934
Db 992 IREELGRIEEQHREKLEVKQLEQALRIODIELKSVTSLNNOVSHTHSENDFHENCML 1051
Qy 935 KKEIAMLKLEVATLAKHOKVKNKYFEDIKILQEKNAELQMTLKLKQKVTGRASQYREQ 994
Db 1052 KKEIAMLKLEVATLAKHOKVKNKYFEDIKILQEKNAELQMTLKLKQKVTGRASQYREQ 1111
Qy 995 LKVLTAENTMLTSKLKE 1011
Db 1112 LKVLTAENTMLTSKLKE 1128
RESULT 9
ABR47548
ID ABR47548 standard; protein; 1341 AA.
XX ABR47548;
AC ABR47548;
XX 12-JUN-2003 (first entry)
DT 12-JUN-2003 (first entry)
XX Breast cancer associated protein sequence SEQ ID NO:334.
DE Breast cancer; cytostatic; gene therapy.
XX Human; breast cancer; cytostatic; gene therapy.
XX Homo sapiens.
OS
XX WO2003004989-A2.
XX
XX
XX 16-JAN-2003.
XX

PF 21-JUN-2002; 2002WO-US019669.
 XX 21-JUN-2001; 2001US-0299887P.
 PR 27-JUN-2001; 2001US-0301572P.
 PR 18-JUL-2001; 2001US-0306501P.
 PR 25-SEP-2001; 2001US-025002P.
 PR 05-MAR-2002; 2002US-0362585P.
 PR 14-MAY-2002; 2002US-0380391P.
 XX (MILL-) MILLENIUM PHARM INC.
 PA Lillie J, Gannavarapu M, Glatt K, Hoerssh S, Kamathkar S;
 PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
 PI Bast RC, Horcobegyi GN, Fuszta L, Meric F, Sahin A, Mills GB;
 XX WPI; 2003-210381/20.
 DR N-PSDB; ACC50246.
 XX Breast cancer diagnosis or treatment by comparing the level of expression
 PT of a marker in a patient sample with that in the control non-breast
 PT cancer sample.
 XX Claim 1; SEQ ID NO 334; 128pp; English.
 XX The present invention describes a method for assessing whether a patient
 CC is afflicted with breast cancer. The method comprises comparing the level
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
 CC ABRA47386 to ABRA47632) in a patient sample and the normal level of
 CC expression of the marker in a control non-breast cancer sample, where a
 CC significant increase in the level of expression of the marker in the
 CC patient sample and the normal level is an indication that the patient is
 CC afflicted with breast cancer. The breast cancer associated sequences from
 CC the present invention have cytostatic activities and can be used in gene
 CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1341 AA;
 SQ
 Query Match 59.4%; Score 3075; DB 6; Length 1341;
 Best Local Similarity 57.4%; Pred. No. 3.8e-199;
 Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;
 QY 1 MVATLLSYGAVIEVQNKASLTPLLLAIQKRSQTVETFLTKNANAFNFSKCTALMLAI 60
 DB 98 VVAKLLSHGAVIEVHNKASLTPLLLSITKRSQIVETFLTKNANAVNKYCTALMLAV 157
 QY 61 CEGSSEIVGMLQQNVVPAEDIHGITAERYAAAGVNYIHQOLLEHTRKLPKQNPNTNP 120
 DB 158 CHGSSEIVGMLQQNVVPAADICGVTAHYAVTCGFHHIEQIMEXIRKLSKNHNTNP 217
 QY 121 EGTSTGTDPDEAAPLA-----ERTPDTAESLLEKTPDEAARLVE 158
 DB 218 EGTSGTDPDEAAPLAERTPDTAESLVEKTPDEAAPLVEKTPDEAASLVE 277
 QY 159 GTSAKIQCLGATSGKFKQSTTEPRKILRPTKTSKFSWP----- 200
 DB 278 GTSDKIQLERATSGKFKQSAETPREITSKETSFTWPAKGRPKRIAWKEDTPR 337
 QY 201 -----AKERSRKITWEEKTSVTECVAGVTPNKTVEKGTSMNIACP 244
 DB 338 BIMPASKETSEKFTWAAKGRPKIAWEKKEPVKTGCVARVTSNKTVEKGRSKMIACP 397
 QY 245 TKETSTKASTNVDSVSPFISLFTGTIENSQCTKVEEDFNLATKIISKAAQNYTCLP 304
 DB 398 TKESSTKASAN----- 408
 QY 305 DATYQDKIKTNHKTEDQMPFSEKREDEEYSDGSLFESSAKTQVCIPESMYQVME 364
 DB 409 -----DQRFSESKQEDDEHYSCDSRLFESSAKIQVCIPESIQVME 452
 QY 365 INREVEELPEKPSAFKPAVEMQKTVPNKAFELKNEQTLRAAQMPPFSESKQKDDDEENSWDS 424

Db 453 INREVEELPEKPSAFKPAIEMQNSVPNKAPELKNQTLRADPMFPPEKQKDEENSWDS 512
 QY 425 EPCETVTSQKDVYLPKATHQKEFDLTSGLKLESVPKDGLLKPTCGRKVSLPNKALELKD 484
 DB 513 ESLCETVTSQKDVCLPKATHQKEIDKINKGLKLESFNDGLLKATCGMKVSIPTKALEL 572
 QY 485 ETFKAESPDGDLKPTCGRKVSLPNKALELKDRETLKAESPDNDGLLKPTCGRKVSLPN 544
 DB 573 QTFKAEPGKPSAFEP-----ATEMQK-----SVN 598
 QY 545 KALELKDRETFKAQMPFSESKQKDDDEENSWDFESFLETLQNDVCLPKATHQKEFDLTS 604
 DB 599 KALELKNQTLRADEILFSESKQKDEENSWDTESLCETVTSQKDVCLPKAAHQKEIDKIN 658
 QY 605 GKLESPPDKDGLLKPTCGMKISLPNKALELKDRETFKAEDVSSVESTSLFKGKPTTNSQ 664
 DB 659 GKLEGSVPKDGLLKANKCMKVSIPTKALELMDMTFKAEPPEK-PSAF-----EPAIEHQK 713
 QY 665 STKVEEDFNLTGEGATKTVTQGOERDGIIERAPQDQTNKMPSTSELGRKEDTKTSDE 724
 DB 714 SV-----PNKALELKNQTLRADEILP-----SESKQKDEEENSWDS 751
 QY 725 IISVSDTQNYCLPEATYQKSIKTNGKLES----- 756
 DB 752 SLCETVTSQKDVCLPKATHQKEIDKINKGLKLESNDGFLKAPCRMKVSIPTKALELMD 811
 QY 757 -----PKPSHFPPATQMNSVPNKGLEWKNKOTLRA----- 788
 DB 812 TFKAEPPEKPSAFEPAIEMQNSVPNKALELKNQTLRADQMPPFSESKQKVEENSWDS 871
 QY 789 -----DSTTLISKILDALPSCBGRRELKKNCEQITA 819
 DB 872 LRETVSQKDVCPKATHQKEMDKISGLKEDSTLSKILDTVHSCERARELQKHCEQRTG 931
 QY 820 KMEQWKKFVQLQKELSEAKIKSOLENQKAKWEOELCSVLPLNQBEEKRNVDILKEK 879
 DB 932 KMEQWKKKFCVLKXKLESAKIKSOLENQKQKWEQELCSVRLTLNQBEEKRNADILNEK 991
 QY 880 IRP-----EEOLRKKLEVKHOLEQTLRIQDIELKSVTSNLNQVSHTHESENDLPHE 934
 DB 992 IREELGRIEEQRKELEVKQOLEQALRIQDIELKSVESNLNQVSHTHESENYLLHENC 1051
 QY 935 KKEIAMLKLEVALTKHQHVKNKYFEDIKLOEKNAELQMTLKLKOKTVTKRSQYREQ 994
 DB 1052 KKEIAMLKLEIATLKHQVQENKVFEDIKILKEKNAELQMTLKLKESLTKRSQYSGQ 1111
 QY 995 LKVLTAENTMLTSKLKE 1011
 DB 1112 LKVLIAENTMLTSKLKE 1128
 RESULT 10
 ADL93227
 ID ADL93227 standard; protein; 1341 AA.
 XX AC ADL93227;
 XX AC ADL93227;
 DT 20-MAY-2004 (first entry)
 XX Human breast cancer-associated polypeptide #49.
 DE gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
 KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
 OS Homo sapiens.
 XX Homo sapiens.
 FN US2003166022-A1.
 XX US2003166022-A1.
 PD 04-SEP-2003.
 XX 04-SEP-2003.
 PF 15-APR-2002; 2002US-00124805.
 XX 15-APR-2002; 2002US-00124805.
 PR 28-DEC-1998; 98US-00222575.
 XX 28-DEC-1998; 98US-00222575.

Tue Feb 22 09:49:15 2005

PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 13-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
PA Houghton RL, Sleath PR, Persing DH;
XX WPI; 2003-874918/81.
XX N-PSDB; ADL93226.
XX An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX Example 12; SEQ ID NO 565; 294pp; English.
XX The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a
CC cancer in a patient and treating a cancer in a patient. The present
CC sequence represents the amino acid sequence of a human breast cancer-
CC associated polypeptide.
XX Sequence 1341 AA;
SQ Query Match 59.4%; Score 3075; DB 7; Length 1341;
Best Local Similarity 57.4%; Pred. No. 3.8e-199; Gaps 12;
Matches 664; Conservative 76; Mismatches 145; Indels 272;
QY 1 MVATLLSYGAVIEVQNKASITPLLLAIQKRSQTVFELTKNANANAFNESKCTALMLAI 60
DB 98 VVAKLLSHGAVIEVHNKASITPLLLSITKRSQTVFELTKNANANAVNKYKCTALMLAV 157
QY 61 CEGSSEIVGMLQONVDVFAEDTHGTAERYAARGVNIHQOLLEHIRKLPKNPONTNP 120
DB 158 CHGSSEIVGMLQONVDVFAADICGVTAHYAVTCVGFHHIHEQIMEYIRKLSKNHONTNP 217
QY 121 EGTSTGTPDEAAPLA-----ERTPDTAESLLEKTPDEAAALVE 158
DB 218 EGTSGTPDEAAPLAERTPDTASLVEKTPDEAAPLVEPTDTAESLVEKTPDEAAALVE 277
QY 159 GTSAKIQCLGKATSGKFEQSTETPRKILPTKETSKEFSWP----- 200
DB 278 GTSKIQCLEKATSGKFEQSAETPREITSPAKETSEKFTWPAKGRPRKIAMEKKEDTPR 337
QY 201 -----AKERSKKTWEKETSVKTECVAGVTPNKTVELEKGTSNMIACP 244
DB 338 EIMSPAKETSEKFTWAAGRPRIKAWKETPVTGCVARTSNKTKVLEKGRSKMIACP 397
QY 245 TKETSTKASTNVDSVSSVEPIESFLGTRTRIENSQCTKVEEDPNLATKIISKSAANYTCLP 304
DB 398 TKESSTKASAN----- 408
QY 305 DATYQKDIKTNHKTEDQMFPSSEKREDEBYSDWSGLFESSAKTOVCIPESMYQKWE 364
DB 409 -----DQRFPSSEKQDEEDYSDRSLSFESSAKIQVCIPESYQKWE 452
QY 365 INREVEELPEKPSAFKPAEMQKTPNKAFAELKNEQTLRAAQMPFSESQKQDEENSDS 424
DB 453 INREVEEPPKPSAFKPALEMONSVNKAFAELKNEQTLRADMPFPFSESQKQDEENSDS 512
QY 425 ESPCETVSQKQYLPKATHQKFDFTLSGKLESVPKDGILLKPTCGRKVSLNPKALELDR 484
DB 513 ESLCETVSQKQVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELDM 572

QY 485 ETFKAESPDKGLLKPTCGRKVSLNPKALELKDRETLKAESPNDGLLKPTCGRKVSLPN 544
DB 573 QTFKAEPFGKPSAFEP-----ATEMCK-----SVPN 598
QY 545 KALELKDRETFRAAQMPFSESQKQDEENSDWPFESFLTLQNDVCLPKATHQKEFDFTLS 604
DB 599 KALELKNEQTRADEILFSESQKQDEENSDWPFESFLTLQNDVCLPKATHQKEIDKIN 658
QY 605 GKLESFPDGLLKPTCGMKISLNPKALELKDRETFKAEDVSSVSTFSLFGKPTTENSQ 664
DB 659 GKLEGSFVKGILLKANCGRKMSIPTKALELMDQMTFKAEPPEK-PSAF-----EPALEMCK 713
QY 665 STKVEEDFNLTKEGATKTVTQOERDGIIFERAPODQTNKMPTELGRKEKDTKSTDS 724
DB 714 SV-----PNKALELKNEQTLRADEILP-----SESKQKQDEESWDSE 751
QY 725 IISVSDTONYECLEPEATYQKEIKTTNGKIEES----- 756
DB 752 SLCETVSQKQVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQ 811
QY 757 -----PEKPSHFEPEATEMONSVNPKGLEWKNKQTLRA----- 788
DB 812 TFKAEPPKPSAFEPALTEMOKSVNPKALELKNEQTLRADQMPFSESQKQKVEENSDSES 871
QY 789 -----DSTTLKILDALPSCERGERELKKNCEQITA 819
DB 872 LRETVSQKQVCLPKATHQKEMDKISGLKEDSTSLSKILDTVHSCERARELQKHCEQRTG 931
QY 820 KMEOMKQKFCVLOKELSEAKEIKSOLENOKAKAEQELCSVRLPLNOBEEKRRNVILKEK 879
DB 932 KMEOMKKKFCVLKKGSEAKEIKSOLENOKQKVEQELCSVRLTLNQEEKRRNADILNEK 991
QY 880 IRP-----EEQLRKLEVKHQLQETLRIOIELKSVTSNINQVSHTHSENDLFHENCML 934
DB 992 IREELGRIEGHRKELEVKQOLEQALAIQDIELKSVESNLNQVSHTHSENYLLHENCML 1051
QY 935 KKEIAMLKLEVALTKHQHVKENKYFEDIKLOKNAELQMTLKLKOKTVTKRASQYREQ 994
DB 1052 KKEIAMLKLEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKESLTKRASQYSGQ 1111
QY 995 LKVLTAENTMLTSKKE 1011
DB 1112 LKVLIAENTMLTSKKE 1128
RESULT 11
ABJ37788
ID ABJ37788 standard; protein; 1349 AA.
XX AC ABJ37788;
XX DT 15-MAY-2003 (first entry)
XX DE Human tumour-related protein - SEQ ID No 573.
XX KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
XX KW tumour; breast cancer; cancer; immune response stimulation.
XX OS Homo sapiens.
XX XX WO200283956-A1.
XX PD 24-OCT-2002.
XX PF 15-APR-2002; 2002WO-US012378.
XX XX 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX PA (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI

PR	22-JUN-2000; 2000US-00604287.
PR	20-JUL-2000; 2000US-00620405.
XX	(CORI-) CORIXA CORP.
XX	Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI	WPI; 2001-611721/70.
DR	N-PSDB; AAS47421.
XX	Breast Tumor Proteins and nucleic acids useful for the prevention,
PT	diagnosis and treatment of breast cancer.
XX	Claim 22; Page 292-295; 297pp; English.
PS	The invention relates to isolated breast tumour proteins and nucleic
CC	acids that encode them, including immunogenic fragments of the proteins.
CC	Also included are expression vectors expressing the proteins, transformed
CC	cells and antibodies raised against the proteins or an antigen presenting
CC	cell expressing the protein. The proteins and nucleic acids may be used
CC	in the prevention, diagnosis and treatment of diseases associated with
CC	inappropriate breast tumour protein expression, i.e. breast tumours and
CC	breast cancer e.g by gene therapy. The nucleic acids and their
CC	complements may also be used as DNA probes in diagnostic assays to detect
CC	and quantitate the presence of similar nucleic acids in samples, and
CC	therefore which patients may be in need of restorative therapy. The
CC	proteins, nucleic acids and antibodies may be used in assays to identify
CC	modulators (e.g. antagonists) of breast tumour protein expression and
CC	activity. The antibodies and antagonists may also be used to down
CC	regulate expression and activity. The antibodies may also be used as
CC	diagnostic agents for detecting the presence of the proteins in samples
CC	(e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC	purification diagnostic techniques. The present sequence is a breast
CC	tumour protein encoded by a cDNA from a breast tumour cDNA library
CC	isolated by subtractive hybridisation against a normal breast CDNA
CC	library
XX	
SQ	Sequence 1095 AA;
Query Match	45.9%; Score 2372.5; DB 4; Length 1095;
Best Local Similarity	52.2%; Pred. No. 1.2e-151;
Matches	529; Conservative 73; Mismatches 151; Indels 261; Gaps 13;
QY	104 LLEHIRKLPKNPQNTPEGTSTGTPDEAAPLAERTPDTAESLLKETPDEARLVEGTSAK 163
DB	24 LLENVISTNPQSVKTE-----YKELLQEFDNDA-----TTKA 58
QY	164 IQCLGKATSGKFEOSTBETPKRI-----LAPTSTSEKFSWPAKERSRK 207
DB	59 IDELKEC---FLNQDTETLSNVFVFMQLIYDSSLCDLFMSPAKETSFKFWAAGRPRK 114
QY	208 ITWEKETSVKTECVAGVTPNKTEVLKGSTNMIACTPKTSTKASTNVDSVSVEPIFSL 267
DB	115 IAWKEKTFTPVTKGVCARVTSNKTVLEKRSKMIACTPKESSTRASAN-----162
QY	268 FGTRTIENSQCTKVEEDFNLATKIISKSAANYTCPLDPATYQDKIKTNHKIEDQMFPSE 327
DB	163 -----DQRFPE 169
QY	328 SKREDEEYSWDGSLPFESSAKTOVCIPESMYQKWEINREVHEELPEKSPAFKPAVEMQK 387
DB	170 SKQEEDESYCDRSLPFESSAKIQVCIESIYQKWEINREVVEEPKKPSAFPAPAEIMQN 229
QY	388 TVPNKAFELKNEQTILRAAQMPFPSESQKDDSENSWSDSSPCBTVSQKDYYLPKATHOKEF 447
DB	230 SVPNKAFELKNEQTILRADPMFPFPSESQKDYBENSWSDESLSCTBVSQKDVCPLPKATHQKEI 289
QY	448 DTLSGLKLES PVKGLLKPTCGRKVYSLPNKALELKDRETFFKAESDPDXGLLKPTCGRKVS 507
DB	290 DRINKGLESPNKGDLKXATCGMKVISPTFKALELKDQMOTTFKAEPPCKGPSAFEPATEMOKS 349
QY	508 LPNKALELKDRETFLKAEPSDNDGLLKPTCGRKVSLPNKALELKDRETFFKAAQMFPSSESQK 567

QY	427	PCETVSKQDVYLKPAATHQKKEFDTLTSGKLKESPVKDGILLKPTCGRKVSLPNKALELKDRET	486
DB	187	LCETVSKQDVCLPKATHQKEIDKINGKLEESPNKDGILLKATCGMKVSIPTKALELKDQMT	246
QY	487	FKAESPDKDGLLKPTCGRKVSLPNKALELKDRETAKESPNDGLLKPTCGRKVSLPNKA	546
DB	247	FKAEPKPKSAFEPATEMOKSVPNKALELKNEQTLRAD-----	284
QY	547	LELKDRETFKAAQMPFSESOKODPEENSWPESFLETLQNDVCLPKATHQKEFDTLTSGK	606
DB	285	-----ETLPSESKQKQYEENSWDTESLCETVSKQDVCLPKAAHQKEIDKINGK	332
QY	607	LEESPDKDGILLKPTCGMKVSLPNKALELKDRETFKAEDVSVESVTFSLFGKPTTENSQST	666
DB	333	LESGPGDGLLKANGCMKVSIPTKALELMDQMTFKABPPEK-PSAF---EPATEMOKSV	387
QY	667	KVEEDFNLTTEGATKVTGQOERDGIIERAPODQTNKMTSELGRKEDTKSTDSEIL	726
DB	388	-----PNKALELKNEQTLRADEILP-----SESKQKQYEESWDSESL	425
QY	727	SVSDTQNYCLPEATYQKEIKTTNGKIBES-----	756
DB	426	CETVSKQDVCLPKAAHQKEIDKINGKLEESPNKDGFLKSPCRMVSIPTKALELMDQMTF	485
QY	757	-----PEKPSHPPEPATEMOKSVNKGLEWKNQOTLRA-----	788
DB	486	KAEPPEKPSAFEPATEMOKSVPNKALELKNEQTLRADQMFPSKQKNVEENSWDSESLR	545
QY	789	-----DSTTLISKILDALPSCERGRELKKNCEQITAKM	821
DB	546	ETVSKQDVCPKATHQKEMDKISGLKEDSTLSKILDTVHSCERARELQKDHCQRTGKM	605
QY	822	EQMKNKFCVLQKELSEAKIISKOLENQAKWQBELCSVRLPLNQOEKKRRNVDIILKEKIR	881
DB	606	EQMKKKFCVLKKLSEAKIISKOLENQKVKWQBELCSVRLTLNQOEKKRRNADIILNEKIR	665
QY	882	P-----EEOLRKKLVGHOLEOTLRIQDIELKSVTSNLNQVSHTHESENDLPHENCMLKK	936
DB	666	ELGRIIEGHRKELEVKQOLEQALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKK	725
QY	937	ETAMLKLEVALTKHQHVKENKYFBDIKILQEKNAELQMTLKLKQKTVTKRASOYREOLK	996
DB	726	ETAMLKLEIATLKHQYOXENKYFEDIKILKEKNAELQMTLKLKXESLTKRASOYSGQLK	785
QY	997	VLTAENTMLTSLKE 1011	
DB	786	VLIAENTMLTSLKE 800	
RESULT 15			
AAU33357			
ID	AAU33357 standard; protein; 1095 AA.		
XX	AAU33357;		
XX	18-DEC-2001 (first entry)		
DT	Human breast cancer protein B726P fusion protein #1.		
XX	Human; ss; breast cancer protein; tumour; cancer; cytostatic;		
KW	gene therapy; immunogen.		
XX	Homo sapiens.		
XX	WO200179286-A2.		
PN	25-OCT-2001.		
XX	12-APR-2001; 2001WO-US012164.		
XX	17-APR-2000; 2000US-00551621.		
PR	08-JUN-2000; 2000US-00590751.		

Db 350 VPNALELEKNEOTLRAD-----EILPSESKQ 375
Qy 568 KDEENSWDFESFLETLQNDVCLPKATHQKEFDTLGKLBESDPDKGLLPTCGMKISL 627
Db 376 KDYEESWDSLSLCETVSQKDVCLPKAXHQEIDKINGKLEGS PVKDGLLKANCCKMYSI 435
Qy 628 PNKALELKDRETFKAEDVSSVESTSLFGKPTTENSQSTKVVEEDPNLTKEGATKTVTGO 687
Db 436 PTKALELMDMTFKAEPPK-PSAF-----EPAIEMQKSV-----PNKALELK 477
Qy 688 QERDIGIITERAPQDQTNKMTPSLGRKEDTKSTDSSELIISVSDTONYECLEPEATYQKEIK 747
Db 478 NEOTLRADAILP-----SESKQDYESSWDSLSLCETVSQKDVCLPKAXHQEID 528
Qy 748 TTNGKIES-----PEKPSHFEPATEMNSV 773
Db 529 KINGLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFPAIEMQKSV 588
Qy 774 PNKGLEWQKQTLRA-----788
Db 589 PNKALELNEOTLRADQMPFSESOKQYVEENSWDSLSLRETYSQKDVCPKATHQKEMDK 648
Qy 789 -----DSTTLISKILDALPSCERGRELKDNCEQITAKMEQMKNFVLOKELSEAKEIK 842
Db 649 ISGKLEDSTLSKILDVTVHSERARELOKCEQRTGMEQMKKFCVLKXKLSAKEIK 708
Qy 843 SOLENQKAKWQELCSVRLPLNQEEKERNVDILKEKIRP-----EEQLRKKLEVHQLE 897
Db 709 SOLENQKVKWQELCSVRLTLNQEEKERNADILNEKIRBELGRIBEQHRKELEVHQLE 768
Qy 898 QTLRIQDIELKSVTSNLNQVSHTHESNDLPHENCMLKKEIAMLKLEVATLKHQHVKEN 957
Db 769 QALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQVQKEN 828
Qy 958 KYFEDIKILQEKNAELOMTLXKOKTVTKRASQYREQLKVLTAENTMLTSKLE 1011
Db 829 KYFEDIKILKERNAELOMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLE 882

Search completed: February 19, 2005, 02:04:50
Job time : 117.137 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:55:34 ; Search time 164.816 Seconds
(without alignments)
4166.462 Million cell updates/sec

Title: US-09-602-362E-23
Perfect score: 6920
Sequence: 1 MTRKKTINLNIDQAQRKA.....NHLKNRIYQVEKEKAETENS 1341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6920	100.0	1341	2	Q9BXX3	Q9bxx3 homo sapien
2	3075	44.4	1011	2	Q9BXX2	Q9bxx2 homo sapien
3	1740	25.1	1710	2	Q9H101	Q9h1q1 homo sapien
4	1736.5	25.1	1715	2	Q9UP58	Q9ups8 homo sapien
5	1718	24.8	424	2	Q9NS19	Q9ns19 homo sapien
6	1246	18.0	1080	2	Q9ZRI4	Q9zr14 homo sapien
7	1112	16.1	823	2	Q9H0H6	Q9h0h6 homo sapien
8	962.5	13.9	1043	2	Q69ZS2	Q69zs2 mus musculu
9	906	13.1	718	2	Q9HCD1	Q9hcd1 homo sapien
10	750	10.8	251	2	Q6PFI4	Q6pfi4 homo sapien
11	745.5	10.8	641	2	Q60311	Q60311 homo sapien
12	736.5	10.6	453	2	Q9IZM7	Q9izm7 homo sapien
13	654	9.5	992	1	AN18_HUMAN	Q9ivf6 homo sapien
14	594	8.6	712	2	Q6S8J5	Q6s8j5 homo sapien
15	570	8.2	544	2	Q6S5H7	Q6s5h7 homo sapien
16	570	8.2	581	2	Q6S5H4	Q6s5h4 homo sapien
17	558.5	8.1	584	1	AN21_HUMAN	Q66yrc homo sapien
18	556	8.0	292	2	Q68DM0	Q68dm0 homo sapien
19	555	8.0	733	2	Q811D2	Q811d2 mus musculu
20	537.5	7.8	546	2	Q9NW69	Q9nw69 homo sapien
21	533	7.7	498	2	Q6S8J7	Q6s8j7 homo sapien
22	532	7.7	264	2	Q9H560	Q9h560 homo sapien
23	522.5	7.6	508	2	Q6S8J2	Q6s8j2 homo sapien
24	520	7.5	486	2	Q6S5H5	Q6s5h5 homo sapien
25	513	7.4	376	2	Q6S8J8	Q6s8j8 homo sapien
26	513	7.4	381	2	Q6NXN7	Q6nxn7 homo sapien
27	507	7.3	394	2	Q6S8J4	Q6s8j4 homo sapien
28	506	7.3	376	2	Q6S8J3	Q6s8j3 homo sapien
29	478	6.9	594	2	Q6S5H6	Q6s5h6 homo sapien
30	474.5	6.9	533	2	Q9CS61	Q9cs61 mus musculu
31	470.5	6.8	1388	2	Q9GL21	Q9gl21 canis famil

ALIGNMENTS

RESULT 1

ID	Q9BXX3	PRELIMINARY;	PRT;	1341	AA.
AC	Q9BXX3				
DT	01-JUN-2001	(TrEMBLrel. 17, Created)			
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)			
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)			
DE	Breast cancer antigen NY-BR-1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21174979; PubMed=11280766;				
RA	Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,				
RA	Jaeger E., Knuth A., Old L.J., Chen Y.T.;				
RT	"Identification of a Tissue-specific Putative Transcription Factor in				
RT	Breast Tissue by Serological Screening of a Breast Cancer Library.,"				
RL	Cancer Res. 61:2055-2061(2001).				
DR	EMBL; AF269087; AAK27325.1;				
DR	GO; GO:0005634; C:nucleus; NAS.				
DR	GO; GO:0005515; F:protein binding; NAS.				
DR	GO; GO:0003700; F:transcription factor activity; NAS.				
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.				
DR	InterPro; IPR002110; ANK.				
DR	InterPro; IPR001969; Pept_Asp_AS.				
DR	Pfam; PF00023; Ank; 6.				
DR	PRINTS; PR01415; ANKYRIN.				
DR	SMART; SM00248; ANK; 6.				
DR	PROSITE; PS50088; ANK_REPEAT; 4.				
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.				
DR	PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.				
KW	ANK repeat.				
SQ	SEQUENCE 1341 AA; 152776 MW; 33853DDE6FD3A58B CRC64;				

Query Match		100.0%;	Score 6920;	DB 2;	Length 1341;
Best Local Similarity		100.0%;	Pred. No. 6e-243;		
Matches 1341;		Conservative	0;	Mismatches	0;
		Indels	0;	Gaps	0;
Qy	1	MTRKKTINLNIDQAQRKATLHWACVNGHEEVVTLVDVRKQQLDVLDEHRTPLMKALQC	60		
Db	1	MTRKKTINLNIDQAQRKATLHWACVNGHEEVVTLVDVRKQQLDVLDEHRTPLMKALQC	60		
Qy	61	HOECANILDSGADINLDVYGNMALHYAVYSILSVAKLLSHGAVIEVHNKASLTPL	120		
Db	61	HOECANILDSGADINLDVYGNMALHYAVYSILSVAKLLSHGAVIEVHNKASLTPL	120		
Qy	121	LLSITKSEQIVFELLIKNANANVVKYKTALMLAVCHGSGSEIVGLLQNVDFVAADI	180		
Db	121	LLSITKSEQIVFELLIKNANANVVKYKTALMLAVCHGSGSEIVGLLQNVDFVAADI	180		
Qy	181	CGVTAHYAVTCGPHHITHEQIMEYIRKLSKNHNTNPEGTSAGTPDEAPLAERTPDPAE	240		
Db	181	CGVTAHYAVTCGPHHITHEQIMEYIRKLSKNHNTNPEGTSAGTPDEAPLAERTPDPAE	240		

181 CGVTAHYAVTCGPHIHEQIMYIRKLSKNHNTNPGTSAGTPDEAPLAERTPDAT 240
241 SLVEKTPDEAPLVERTPDABSLVEKTPDEASLVEGTSKIQCLEKATSGFEQSAE 300
241 SLVEKTPDEAPLVERTPDABSLVEKTPDEASLVEGTSKIQCLEKATSGFEQSAE 300
301 TPRTITPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
301 TPRTITPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
361 IAWKKEPTVTKGVARVTSNKTIVLEKGRSKMTACPTKESSTKASANDORFSESQEE 420
361 IAWKKEPTVTKGVARVTSNKTIVLEKGRSKMTACPTKESSTKASANDORFSESQEE 420
421 DEEYSCDSRSIFESSAKIQVICPSIYQKWEINREVEEPPKPSAFKPALEIMQNSVPNK 480
421 DEEYSCDSRSIFESSAKIQVICPSIYQKWEINREVEEPPKPSAFKPALEIMQNSVPNK 480
481 AFEKNEOTLRADPMFPPEKQDYENSWSSESLCETVSQKDVCLPKATHQKIDKING 540
481 AFEKNEOTLRADPMFPPEKQDYENSWSSESLCETVSQKDVCLPKATHQKIDKING 540
541 KLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKABPPKPSAFBEPATEMOKSVNKA 600
541 KLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKABPPKPSAFBEPATEMOKSVNKA 600
601 LELKNEQWRADEIILPSESQKDYENSWSSESLCETVSQKDVCLPKAAHQKEIDKINGK 660
601 LELKNEQWRADEIILPSESQKDYENSWSSESLCETVSQKDVCLPKAAHQKEIDKINGK 660
661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKABPPKPSAFBEPATEMOKSVNKA 720
661 LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKABPPKPSAFBEPATEMOKSVNKA 720
721 ELKNEQWRADEIILPSESQKDYENSWSSESLCETVSQKDVCLPKATHQKIDKINGKL 780
721 ELKNEQWRADEIILPSESQKDYENSWSSESLCETVSQKDVCLPKATHQKIDKINGKL 780
781 EESPDNDGFLKAPCRMKVISIPTKALELMDQTFKABPPKPSAFBEPATEMOKSVNKALE 840
781 EESPDNDGFLKAPCRMKVISIPTKALELMDQTFKABPPKPSAFBEPATEMOKSVNKALE 840
841 LKNEQWRADEIILPSESQKDYENSWSSESLCETVSQKDVCLPKATHQKIDKINGKL 900
841 LKNEQWRADEIILPSESQKDYENSWSSESLCETVSQKDVCLPKATHQKIDKINGKL 900
901 DSTLSKILDTVHSCERARELOKHCEQRTGKMEQMKKFCVLKXKLSPAKIKSLEQ 960
901 DSTLSKILDTVHSCERARELOKHCEQRTGKMEQMKKFCVLKXKLSPAKIKSLEQ 960
961 KVKWEQELCSVRLTLNQEERKNADILNEKIPREELGRIEEQRKLEVKQOLEQALRIQ 1020
961 KVKWEQELCSVRLTLNQEERKNADILNEKIPREELGRIEEQRKLEVKQOLEQALRIQ 1020
1021 DIEKSVESNLQVSHTHENENYLLHNCMLKEIAMLKLEIATLKHQYQEKENKYFEDI 1080
1021 DIEKSVESNLQVSHTHENENYLLHNCMLKEIAMLKLEIATLKHQYQEKENKYFEDI 1080
1081 KILKEKNAELQWTLKLEESLTKRASQYSGQLKVLIAENTMWSKJKEKQDKLEIAEIE 1140
1081 KILKEKNAELQWTLKLEESLTKRASQYSGQLKVLIAENTMWSKJKEKQDKLEIAEIE 1140
1141 SHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQKKNVDVSSITLYNNEVLHPLSEA 1200
1141 SHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQKKNVDVSSITLYNNEVLHPLSEA 1200
1201 ORKSKSLKINLYAGDALRENTLVSEHAQORDQTOCQKMEABEHYQNEQDNVVKHTEQQ 1260
1201 ORKSKSLKINLYAGDALRENTLVSEHAQORDQTOCQKMEABEHYQNEQDNVVKHTEQQ 1260
1261 ESLDQKLFQLOSKNWLQQQLVHAHKKADNKSKITIDIHFLERKQVHLLKERNEEIFNY 1320
1261 ESLDQKLFQLOSKNWLQQQLVHAHKKADNKSKITIDIHFLERKQVHLLKERNEEIFNY 1320

1321 NNHLKNRIIYOYEKEKAETENS 1341
1321 NNHLKNRIIYOYEKEKAETENS 1341
RESULT 2
Q9BX2 Q9BX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BX2;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutharia; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karch J.,
RA Jaeger D., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a tissue-specific putative transcription factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269088; AAK27326.1; --
DR HSP; O75832; LUOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00088; ANK REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
FT NON_TER 1011 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DB438 CRC64;
Query Match 44.4%; Score 3075; DB 2; Length 1011;
Best Local Similarity 57.4%; Pred. No. 8.3e-104;
Matches 664; Conservative 76; Mismatches 145; Indels 272; Gaps 12;
QY 98 VVAKLLSHGAVIEVHNKASLTPLLSITKSEQIVFLLIKNANANVKNKYKCTALMLAV 157
Db 1 MVATLLSGAVIEVQNKASLTPLLSITKSEQIVFLLIKNANANVKNKYKCTALMLAI 60
QY 158 CHGSEIYVGMLLQONVDVFAADICGVTAHYAVTCGPHIHEQIMYIRKLSKNHNTN 217
Db 61 CEGSEIYVGMLLQONVDVFAEDIHGITAERYAAARGVNYIHQQLLEHIRKLPKPNQNTN 120
QY 218 EGTSGTGTDEAPLAERTPDABSLVEKTPDEAPLVERTPDABSLVEKTPDEAPLVE 277
Db 121 EGTSGTGTDEAPLAERTPDABSLVEKTPDEAPLVERTPDABSLVEKTPDEAPLVE 158
QY 278 GTSKDIQCLEKATSGFEQSAEETPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPR 337
Db 159 GTSKDIQCLEKATSGFEQSAEETPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPR 200
QY 338 EIMSPAKETSEKFTWPAKGRPRKIAWEKEDTPRVTGCVARVTSNKTIVLEKGRSKMTAC 397
Db 201 -----AKERSRKITWEEKTSVKTECVAGVTPNKTEVLEKTSNNIACP 244
QY 398 TRESSTKASAN----- 408
Db 245 TRESSTKASTNVDVSSVEPIFSLFGTRTITENSQCTKVEEDFNATKILSKSAQNYTCLP 304
QY 409 -----DQRFPSSEKQEEDEEYSCDSRSIFESSAKIQVICPSIYQKWE 452
Db 305 DATYQDKITNKIEDQMFPSSEKQEEDEEYSCDSRSIFESSAKIQVICPSIYQKWE 364
QY 453 INREVEEPPKPSAFKPALEIMQNSVPNKAFELKNEOTLRADPMFPPEKQDYENSWS 512
Db 365 INREVEEPPKPSAFKPALEIMQNSVPNKAFELKNEOTLRADPMFPPEKQDYENSWS 424

```
Qy 513 ESLCETVSQDVCLPKATHQKEIDKINGKLBESPNVQGLLKATCGMKVSIPTKALELMD 572
Db 425 ESPCETVSQDVCLPKATHQKEIDKINGKLBESPNVQGLLKATCGMKVSIPTKALELMD 484
Qy 573 QTFKAEPGKPSAFEP-----ATEMQK-----SVPN 598
Db 485 ETFKAEPDGLKPTCGRKVSIPTKALELMDKRETLKAESPNDGLLKPTCGRKVSIPTN 544
Qy 599 KALELKNQWRADLILPESKQDYENSWDTSCLCETVSQDVCLPKAAHQKEIDKIN 658
Db 545 KALELKNQWRADLILPESKQDYENSWDTSCLCETVSQDVCLPKAAHQKEIDKIN 604
Qy 659 KLESGPVKQGLLKANGCMKVISPTKALELMDQTFKAEPPEK-PSAF-----EPAIEMQK 713
Db 605 KLESGPVKQGLLKANGCMKVISPTKALELMDQTFKAEPPEK-PSAF-----EPAIEMQK 664
Qy 714 SV-----PNKALELKNQWRADLILP-----SESKQDYENSWDSE 751
Db 665 STKVEEDPNLTKEGATVTVGQERDGIITRAPQDQTNQMTSELGRKEDTTSQDSE 724
Qy 752 SLCETVSQDVCLPKATHQKEIDKINGKLBESPNVQGLLKATCGMKVSIPTKALELMDQ 811
Db 725 IISVSDIQNYECLPEATVQKEIKTTNGKIEES-----756
Qy 812 TFKAEPEKPSAFEPATTEMQSVNPKALELKNQWRADLILPESKQDYENSWDSE 871
Db 757 -----PEKSHFEPATTEMQSVNPKALELKNQWRADLILPESKQDYENSWDSE 788
Qy 872 LRETVSQDVCLPKATHQKEIDKINGKLBESPNVQGLLKATCGMKVSIPTKALELMDQ 931
Db 789 -----DSTLUSKILDALPSCERGRELKKNQCEQITA 819
Qy 932 KMEQMKKFCVLKKESEAKIKSQLENQKQVQELCSVRLTLNQBEKERNADIINEK 991
Db 820 KMEQMKKFCVLKKESEAKIKSQLENQKQVQELCSVRLTLNQBEKERNADIINEK 879
Qy 992 IREELGRIEEQRKELEVKQLEALRIQDIELKSVESNLNQVSHENNYLHENCML 1051
Db 880 IRP-----EQURKKEVQHQLEQTLRIQDIELKSVESNLNQVSHENNYLHENCML 934
Qy 1052 KKEAMUKLEATLKHQVQENKVFEDIKILKEKNAELQMTLKEESLTKRASQYSGQ 1111
Db 935 KKEAMUKLEATLKHQVQENKVFEDIKILKEKNAELQMTLKEESLTKRASQYSGQ 994
Qy 1112 LKVLIAENTMLTSKLKE 1128
Db 995 LKVLIAENTMLTSKLKE 1011

RESULT 3
Q9H1Q1 PRELIMINARY; PRT; 1710 AA.
AC Q9H1Q1
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BA145E8.1 (KIAA1074).
GN Name=BA145E8.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR HSSP; P20749; 1k1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS00086; ANK_REPEAT; 4.
```

```
DR PROSITE; PS00297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE 1710 AA; 196409 MW; 01CBF9BADB894872 CRC64;

Query Match 25.1%; Score 1740; DB 2; Length 1710;
Best Local Similarity 33.8%; Pred. No. 3.2e-55;
Matches 479; Conservative 204; Mismatches 376; Indels 358; Gaps 34;

Qy 10 LNIQDAQKRTALHWACVNGHEEVVTLVDKRCQLDVLGDEHRTPLMKALQCHQECACATIL 69
Db 73 LNDKQKRTALHLACANGHEEVVTLVDKRCQLNCDNENRTALMKAVQCEKCATIL 132
Qy 70 IDSGADINLVGYGNMHLHAYVSEILSVAKLJSHGAVIEVHNKASITPLLSITKXSE 129
Db 133 LEHGADPNLADVHGNTALHYAVYNEIDISVATKLLLYDANIEAKNKDDLTPLLLAVSGKKQ 192
Qy 130 QIVGFLLIKNANANVKNKCTALMLAVCHSGSSEIVGMLLQNVNDVFAADICGVTAHVA 189
Db 193 QMVFEFLIKKANVNAVVKLESS-----214
Qy 190 VTCGFPHIHEQIMFYIRKLSKNHQNTPGTSAGTPDEAAPLAERTPDTAESLVEKTPDE 249
Db 215 -----HQLISYKEERIPKHSQNSNSVDESEDSLSRLSGK-PGV-----DD 256
Qy 250 AAPLIVERTPDTAESLVEKTPDEAAASLVGTSKIQCLEKATSGKPEQSAEETPR--EITS 307
Db 257 SWP-----TSDD-----EDLNFDTKQVPKPSLAKLMTASQOSRKNLEATY 296
Qy 308 PAKETSEKFTWPAKGRPRKIAWEKEDTPRIMSPAKETSEKFTWAAKGRPRKIAWEKKE 367
Db 297 GTVRTGN-----RTLFEDRDSQDEVVVESELT-----325
Qy 368 TPVKTGCVARVTSNKTIVLEKGRSMIACP--TKESSTKASANDQRFPSSEKQEDDEYS 425
Db 326 TSIRKVCFSHTYQSPDLLPKPSHSLANPGLMKEETK-----PGIAKENGID- 376
Qy 426 CDSRSLFESSAKIQVCIPESYQKVMENREVEBPCKPSAFKPAIEHQNSVNPNAFELK 485
Db 377 -----IESAP-----LEQTN--DNLTYYVDE 395
Qy 486 NEQTLRADPMPFPSPKQDYENSWDSESLCETVSQDV-CLPKATHQKEIDKINGKLEE 544
Db 396 VHKNNRSDMSALGLGQEEEDIESPWSSESINFPQKYVDPLAGAADGCKENIGNEQAE- 454
Qy 545 SPNKDGLLKATCGMKVSIPTKALELKD-----MOTFKAEPPKPSAFAPATTEMQSVNPK 599
Db 455 -----DVFIYFSC-MSGSENFMAKLEDTRNVMVAHMESEBYVILHKTTEMKDSVNP 509
Qy 600 ALELKNQWRAD-----EILPESKQDYENSWDTSCLCETVSQDVCLPKAAHQKEID 655
Db 510 AGMKDVQTSKAAEHDLVASEEEREGSEN-----NQPVVEERKCHRNEM 558
Qy 656 KINGKL-----EGSPVKDGLLKANGCMKVISPTKALELMDQTFKAEPPEKPSAFEP 707
Db 559 EVSANIHDGATDDAEDDDDDGLIQKR-----KSGE-TDHOQFPKRNKEYVASSGP 608
Qy 708 AIEMQSVNPKALELKNQWRADLILPESKQDYENSWDTSCLCETVSQDVCLPKAAHQKEID 767
Db 609 ALQMK-----EVKTEKEKTSKES-----628
Qy 768 THQKEIDKINGKLBESPNVQGLLKATCGMKVSIPTKALELMDQTFKAEPPE- --KPSAF 824
Db 629 -----VNSPVFGKASLLTGGLLQVDDSSLSLSEIDEDEGRPT-- 664
Qy 825 EPAIEMQSVNPKALELKNQWRADLILPESKQDYENSWDSESLCETVSQDVCLPKAAHQKEID 884
Db 665 -----KKTSEKKNVKNQIQSMDD-----VDDLTQSSETASE-----DCSLP 701
Qy 885 KATHQKEM---DKISGKLEDSTLSKILDTVHSCERARELOKDHCEQRTKMEOMKKKFC 941
Db 702 HSSYNNFMILLIEQLGMECKDSVLLKQDAALSCEKRLLEKKNHCELLTVKIKMEDKVN 761
Qy 942 VLKXKJSEAKEIKSQLENQKQVQELCSVRLTLNQBEKERNADIINEKIREELGRIEE 1001
```



```

QY 1076 YFEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTWMLTSLKJL-EKQDKET 1134
DB 947 CFEDLKIVKKEQEDLQKTIKONEETLTQISQYNGRLSVLTAENAMLSKLENEKQSKER 1006
QY 1135 LEAETESHPRILASAVODHDQIVTSRKSQEPAPHIAGDAC--LQKNNVDVSTIYNNEV 1192
DB 1007 LEAEVESHRLAALAHDRQDQSETSKRELELAFQARDECSRLQDKQNFVSNLKNDDNEI 1066
QY 1193 LHQPLSEARQSKSLKINLYAGDALRENTLVSEHAQRDQRETQCCQKAEAHMYQNEQDN 1252
DB 1067 LSQQLFKTESKLSLEIEFHHTRDALRENTGLGERVQKDLSTQCCQKMEQKYQNEQV 1126
QY 1253 VNKHTEQBSLQKFLQSQKNNWMLQOQLVHAHKADNKSITIDI----HFLERKQHH 1308
DB 1127 VNKYIGQBSVEERLSQSENNLFRQQLDDAHNKADNKEKTVINIQDQFHAIVQKQAE 1186
QY 1309 -----LLKEKNEEIPYNNHLKNRIYQYEKEAETE 1339
DB 1187 SEKQSLLEERNKELISECNHLKEROYQYENKAE 1223

RESULT 5
Q9NSI9
ID Q9NSI9 PRELIMINARY; PRT; 424 AA.
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DE PREDA protein (Fragment).
GN Name=PREDA;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,
RA Toyoda A., Ishii K., Totoki Y., Choi D.K., Soeda E., Ohki M.,
RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudon J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis P.,
RA Leinhardt R., Raspo M.L.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL163203; CAB90394.1; -.
FT NON TER 1
FT NON TER 424
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

Query Match 24.8%; Score 1718; DB 2; Length 424;
Best Local Similarity 79.7%; Pred. No. 4.8e-55;
Matches 350; Conservative 30; Mismatches 43; Indels 16; Gaps 4;

QY 901 DSTLSKILDTVHSCERARELOKDHCEQRTGQEQKMKKFCVLKYLSEAKIKSQLENQ 960
DB 2 DSTLSKILDTVHSCERARELOKDHCEQRTGQEQKMKKFCVLKYLSEAKIKSQLENQ 61
QY 961 KVKVEQELCSVRLTNQBEERKNADILNEKIREELGRREEQHRKEVKKQLEALRIQ 1020
DB 62 KAKVEQELCS-----KKRDVDILKEKIRP-----EQLRKKLEVKQLEALRIQ 107
QY 1021 DIELKSVESNLNQVSHTHENNYLHENCMLKKEIAMLKLEIATLKQYQENKRYFEDI 1080
DB 108 DIELKSVESNLNQVSHTHESNDLFHENCMLKKEIAMLKLEIATLKQYQENKRYFEDI 167
QY 1081 KILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTWMLTSLKJL-EKQDKET 1140

```

```

DB 168 KILQEKNAELQMTLKLKQKTLTKRASQYREQQLKVLTAENTWMLTSLKJL-EKQDKET 227
QY 1141 SHHPLASAVQDHDQIVTSRKSQEPAPHIAGDACLQKNNVDVSTIYNNEVHLQPLSEA 1200
DB 228 SHHPLASALQDHDQIVTSRKSQEPAPHIAGDACLQKNNVDVSTIYNNEVHLQPLSEA 287
QY 1201 QKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCCQKAEAHMYQNEQDNVKNKTEQ 1260
DB 288 QKSKSPKINLYAGDDLRENALVSEHAQRDQRETQCCQKAEAHMYQNEQDNVKNKTEQ 347
QY 1261 ESLDQKLFLQSQKNNWMLQOQLVHAHKADNKSITIDIHFLERKQHHLLKEKNEEIPY 1320
DB 348 ESLQKLFLKESKNNWMLQOQLVHAHKADNKSITIDIHFLERKQHHLLKEKNEEIPY 405
QY 1321 NNHLKNRIYQYEKEAETE 1339
DB 406 GNHLKECIDQYEKEAETE 424

RESULT 6
Q6ZR14
ID Q6ZR14 PRELIMINARY; PRT; 1080 AA.
AC Q6ZR14;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE Hypothetical protein FLJ46736.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Trachea;
RA Kaneshiro K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RA Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK128577; BAC87508.1; -.
FT NON TER 1080
SQ SEQUENCE 1080 AA; 126741 MW; 47BBA9A933AE595C CRC64;

Query Match 18.0%; Score 1246; DB 2; Length 1080;
Best Local Similarity 41.2%; Pred. No. 1.6e-37;
Matches 307; Conservative 122; Mismatches 200; Indels 116; Gaps 16;

QY 678 KVSIPTKALELMDMOTFKAE-----PPEKPSAPEPAIMQKSVNKALELKN---EQLTRA 730
DB 64 KVSLEELQ--DMQRFNKGMLKVEFQALEKEKVKQLEVEERKKNRNMESANI 121
QY 731 DEILPSESQKQDYESSWDSLSCTVTSQKDVCLPK-----ATHQEKIDKNG 779
DB 122 HDGATDDAEDDDDDGLQKRGSGTDHQ---PFRKENKEYASSGPAALQMKVEKSTKE 178
QY 780 LEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPE---KPSAPEPAIMQKSVN 836
DB 179 KRTSKES---VNSPVFGKASLLTGLQLVDDSSLSLSEIDEDSGRPT-----KKTSN 226
QY 837 KALELKNQTLRADQMPFSESQKQKVEENSWSLSRETSTVTSQKDVCPKATHQKEM---D 893
DB 227 EKNVKNQIQSMDD-----VDLTQSSETASE-----DCELPSSVKNFALLIE 270
QY 894 KISGKLEPSTLSKILDTVHSCERARELOKDHCEQRTGQEQKMKKFCVLKYLSEAKI 953
DB 271 QLGHECKDSVLLKIQDAALSCERLLELKNKNCCELLTVKIKKMDKVNVLQELSETKEI 330
QY 954 KSOLENQKVKVEQELCSVRLTNQBEERKNADILNEKIREELGRREEQHRKEVKKQLE 1013
DB 331 KSOLEHQKVEERELCSLRFSLNQBEERKNADTLTYEKIREQLRRKESQYRKEVEVKQOL 390

```


Db 296 KBRNVEELHOKVREKLITEEQRIEADVTPIKPAKLSAEVLTKGNNNSQVSETDE 355
 QY 1040 NENYLLHNCMLKKEIAMLKLEIATLKHOYQYQKKNYFEDIKILKKNALOMTLKKEE 1099
 Db 356 KED-LLHENRLMQDEIARLEKDTIKQNLK--KYLKDFIVVRKHEDLQKALKRNGE 412
 QY 1100 SITKRASQVSGOLKVLIAENTMLTSL-KEKODKILAEATESHPRLASAVODHDQIVT 1158
 Db 413 TLIAKTIACYSQGLAALTDTNTLRSLKLEKQRESRORLETEMOSYICRLNARCDHQSHS 472
 QY 1159 SRKSQEPAPHIAGDAC--LQRKWNVDVSTIYNNEVLHQPLSEAQKSKSLKININAYAGD 1216
 Db 473 SKEDQELAFQGVDRCHLQENLHVL-----ILSLSKASKSRVLKTELHYTGE 525
 QY 1217 ALRENTLVSEHAQRORETOCQKAEHMYQNEQNVNKHTEQOESLQOKLPOLSKNMW 1276
 Db 526 ALKERALVFEHVQSELKQSQMKDIEKMYKSYNTMEKICBQK-----RFOQLKQKNL 581
 QY 1277 LOOQLVHAHKADNKSKITIDTHF-----LERKMOHH--LLKKEKEIEFNYNHKN 1326
 Db 582 LOOQLDDANKADNQKALINIQARCDARVQNLQAECKRHLLEEDKQVLNETHSK 641
 QY 1327 RIYQYKEKEAETE 1339
 Db 642 KECQYKEKEARE 654

RESULT 10

Q6PF14 PRELIMINARY; PRT; 251 AA.
 AC O6PF14; 27, Created
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toohyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX Strausberg R.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC057764; AAH57764.1; -.
 DR HSP; P20749; IKLB.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR001969; Pept_Asp_AS.
 DR Pfam; PF00023; Ank; 5.

DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 5.
 DR PROSITE; PS0088; ANK_REPEAT; 3.
 DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 DR PROSITE; PS0141; ASP_PROTEASE; UNKNOWN_1.
 KW ANK repeat; Hypothetical protein.
 SQ SEQUENCE 251 AA; 27543 MW; 85934A3579801A3A CRC64;
 Query Match 10.8%; Score 750; DB 2; Length 251;
 Best Local Similarity 77.1%; Pred. No. 3.3e-20;
 Matches 148; Conservative 19; Mismatches 25; Indels 0; Gaps 0;
 QY 4 RKKTINLNIODAQRTALHWAACVNGHEVTVFLVDRKQDLVDGEGHRTPLMKALQCHOE 63
 Db 59 KKTMDLNIRDAKERTALYACANGHAEVTVLLVDRKQDLVDGEGHRTPLMKALQCHOE 118
 QY 64 ACANILDSGADINLVYGNMALHYAVYSEILSVAKLLSHGAVIEVHNKASLTPLLS 123
 Db 119 ACANILDSGADPNVDVYGNVAVNSVAVKLLSCGADIEVKNKAGHTPLLLA 178
 QY 124 ITKSEQIVFLLIKNANANVYKCTALMLAVCHGSSEIVGMLLQNVDFVFAADICGV 183
 Db 179 IRKSEIEVFELTKNANANAVDFKCTALMLAICHGSSEIVGKLLQNVDFVFAADICGV 238
 QY 184 TAEHYAVTCGFH 195
 Db 239 IAERYAVACGFN 250
 RESULT 11
 O60311 PRELIMINARY; PRT; 641 AA.
 AC O60311; 07, Created
 DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
 DE KIAA0565 protein (Fragment).
 GN Name=KIAA0565;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N.,
 RA Chara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:31-39(1998).
 DR EMBL; AB011137; BAA25491.2; -.
 FT NON_TER 1
 SQ SEQUENCE 641 AA; 75777 MW; F11566F2E6A0B9F4 CRC64;
 Query Match 10.8%; Score 745.5; DB 2; Length 641;
 Best Local Similarity 42.0%; Pred. No. 1.2e-19;
 Matches 179; Conservative 69; Mismatches 111; Indels 67; Gaps 6;
 QY 972 RUTLQNEEKKRNADILNEKIEELGRTEOHRKEVQKLEQALRIODIELKSVESNL 1031
 Db 6 RLAKQNEEKKRNADMLYNKDSQLRIKEEBCGVVETKQQLKWNLRVLKELRTVRNL 65
 QY 1032 -----NOVSTHTEHYVLLHEN 1048
 Db 66 DLVQVRNDAQQLSEQDARILQDILTSKQKELMARKKNVSESHRQKEDLFHED 125
 QY 1049 CMLKKEIAMLKLEIATLKHOYQYQKKNYFEDIKILKKNALOMTLKKEESLTKRASQY 1108
 Db 126 CMLQEBIALRLLEIDITKQNKQKKEKYPEDIEAVKKNNDLQKIKLNEETLTILQY 185
 QY 1109 SGOLKVLIAENTMLTSKLKE-KQDKLEILAEIESHHPRLASAVQDHDQIVTSK ----- 1160

```
Db 186 SQGLNLTAEKILNSELNGKQNERLBIENESYRCRLAAAVRCDQSQTLARDLKLDFQ 245
Qy 1161 KSOEPAFHIAQDACLQKQKNDVSVTYNNVHOPLSAQRKSKSLKINLVAGDALRE 1220
Db 246 RTRQEWVR-----LHDKMKVDMGLOAKVEILSEKLSNAESKINSLOIQLHTRDLCGR 299
Qy 1221 NTLVSEHAQRDQRETQCKQKEAHMYQNEQDNVNNKHTEQOESLDQKLFQLOSKNNWLOQQ 1280
Db 300 ESLILERVORDLSQTCQCKKETEQMYQIBQSLKKYIAKQESVEERLSQLOQENMLLRQ 359
Qy 1281 LVHAKKADNKSII--TIDIHF-----LKKMQHLLKKEKNEIFNYYNNHLKNRIVQ 1330
Db 360 LDDAHKANSQKSTSTIQDPHSAKQLQAESEKQILSQLOKNEKMDLDEYNHLKERMDQ 419
Qy 1331 YEKEKA 1336
Db 420 CEKEKA 425

RESULT 12
Q81ZM7
ID Q81ZM7 PRELIMINARY; PRT; 453 AA.
AC Q81ZM7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Melanoma-associated antigen (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bruno R., d'Orlando O., Altomonte A., Lamaj E., Maio M., Pucillo C.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF543495; AAN40505.1; -
FT NON TER 1
SQ SEQUENCE 453 AA; 53140 MW; 81DF80012122BBCA CRC64;

Query Match 10.6%; Score 736.5; DB 2; Length 453;
Best Local Similarity 42.8%; Pred. No. 1.8e-19;
Matches 172; Conservative 83; Mismatches 120; Indels 27; Gaps 8;

Qy 951 KEIKSQLENQKVKWEQELCSVRLTLNQBEEKRNADILNEKIEBELGRIEEOHRKELEVY 1010
Db 2 BEIKSQLKHEILELEKELCSLRFATIQEKKKRNVEEVHQVREKLRITEEQYRIEADVT 61
Qy 1011 QOELQALRIQDIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYQ 1070
Db 62 KPDKPKAKSAEVELKTGGNSNQVSETDEKED-LLHENRLMQDEIARLREKDTIKNQWL 120
Qy 1071 EKENKYFEDIKILKEKNAELQMTKLKBSLTKRASQVSGQLKVLIENMTSLK-KEK 1129
Db 121 EK--LYLKDFEIVKRGKHEDLQALKRNGETLAKTIACVSGQAULTDENTILRSKLEQR 178
Qy 1130 QDKTELEAIEHSHPRLASVDHDIQVTSRKSQEPAPHIAGDAC--LQKKNVDVSVTI 1187
Db 179 ESRQLETEMQSHYCHRLNARCDHQSHSSKKRQDLAFQGTVDKCRHLQENLNSHVL--- 235
Qy 1188 YNNEVLHOPLSAQRKSKSLKINLVAGDALRENTLVSEHAQRDQRETQCKQKEAHMYQ 1247
Db 236 ---ILSLQLSKAESKSRVLKTELHYTGEALKEKALVPEHVQSELSQKQSQMKDIEKMYK 291
Qy 1248 NEQDNVNNKHTEQOESLDQKLFQLOSKNNWLOQLVHAKKADNKSIIITIDIHF----- 1300
Db 292 SCYNTMEXCIEKQE----RFOQLKQKNMLLOQLDDARNKADNQSKAILNIQARCDARVQ 347
Qy 1301 -LERKMQH--LLKEKNEIFNYYNNHLKNRIVQYEKEKAETE 1339
Db 348 NLQASCRKRLLEEDNKMVLNMLNHSKEKECQYEKEKAERE 389

RESULT 13
AN18 HUMAN
ID AN18 HUMAN STANDARD; PRT; 992 AA.
AC Q81VF6; Q72468;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE Ankyrin repeat domain protein 18A.
GN Name=ANKRD18A; Synonym=KIAA2015;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nagase T., Kikuno R., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-349 FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB095935; BAC23111.1; ALT_INIT.
DR EMBL; BC056266; AAH56266.1; ALT_TERM.
DR Genew; HGNC:23643; ANKRD18A.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS00297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS0088; ANK_REPEAT; 4.
KW ANK repeat; Repeat.
FT REPEAT 67 96 ANK 1.
FT REPEAT 100 129 ANK 2.
FT REPEAT 133 162 ANK 3.
FT REPEAT 166 195 ANK 4.
FT REPEAT 199 228 ANK 5.
FT CONFLICT 130 130 K -> E (in Ref. 2).
FT CONFLICT 288 288 Missing (in Ref. 2).
SQ SEQUENCE 992 AA; 115664 MW; B9B288F087340D9A CRC64;

Query Match 9.5%; Score 654; DB 1; Length 992;
Best Local Similarity 18.2%; Pred. No. 3.9e-16;
```

Matches 244; Conservative 121; Mismatches 197; Indels 780; Gaps 13;	
Qy	1 MTKRKTNTLNQDAQKATLHACVNGHEEVVTFVDRKQQLDVLGDEHRTPLMKALQC 60
Db	54 LTRFR--DLARDKRTVLHLACAHGRVQVVTLLHRRQCQIDCDRLNRTPLMKAVHS 111
Qy	61 HOEACANILISGADINLVVGNMALHYAVTSEILSVVAKLLSHGAVIEVHNKASUTPL 120
Db	112 QBEACAIVLLECCANPNKIDYGNALHYAVNKGTSIAERLLSHANIEALNKGNTPL 171
Qy	121 LLSITKRSQIVFELLKNNANAVNKYCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180
Db	172 LFAINSRQHMVEFLKNOANIHADVNFRTALILAVOHNLSIVTLLQONIRISSQDM 231
Qy	181 CGVTAHYAVTCGPHHIEQIMEYIRKLSKNHNTNPEGSTAGTDEAAPLAERTPDAT 240
Db	232 FQCTAEDYALCSDLSIRQILHKNKMLKNHLRNDNQETA-- 273
Qy	241 SLVEKTPDEAALVTPDPAESLVEKTPDEAASLVEGTSKIQCLEKATSGKPEQSAE 300
Db	274 ----- 273
Qy	301 TPREITSPAKETSEKFTWPAKGRPKRIAWKEDTPREIMSPAKETSEKFTWAAKGRPK 360
Db	274 -----MKPA----- 277
Qy	361 IAWEKETPVKGCVARVTSNKTVLKGRSOMIACPTKESSTKASANDORPPSEKQEE 420
Db	278 -----NLKGRERAKAE-----HNLKVASEKQEE- 301
Qy	421 DBEYSCDSRLPSSAKIQVCIPESIQYKVMENREVEBPCKPSAFKPAIEMQNSVFNK 480
Db	302 ----- 301
Qy	481 APELKNQTLRADMPPEPKQDYENSWSDESCEVTSQKDVCLPKATHQKEIDKING 540
Db	302 ----- 301
Qy	541 KLEESNKGGLLKATCGMKVSIPTKALELDMQTFKAEPPGKPSAFEPATEMOKSVPNKA 600
Db	302 RLQRSENKQ----- 310
Qy	601 LELKNQETRADEILPSESKQDYENSWSDESCEVTSQKDVCLPKAAHQKEIDKING 660
Db	311 -----FQDSQ----- 315
Qy	661 LEGSPVKDGLLKANCCKVSIPTKALELMDMOTFKAEPEKPSAFEPATEMOKSVPNKAL 720
Db	316 ----- 315
Qy	721 ELKNEQTLRADEILPSESKQDYENSWSDESCEVTSQKDVCLPKATHQKEIDKINGKL 780
Db	316 -----SYGKKD----- 322
Qy	781 EESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPEKPSAFEPATEMOKSVPNKALE 840
Db	323 ----- 322
Qy	841 LKNEQTLRADMPPEPKQDYENSWSDESCEVTSQKDVCLPKATHQKEMDKISGKLE 900
Db	323 ----- 322
Qy	901 DSTSLSKILTVHSCERARELQDHCERTGKMEQMKKFCVLKCKLSEAKEIKSQLENQ 960
Db	323 ----- 322
Qy	961 KVKWEQELCSVRLTNQOEKRRNADILNEKIRELGRIEQHRKELEVKQLEQALRIQ 1020
Db	323 ----- 322
Qy	1021 DIEKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYPEDI 1080
Db	323 -----AMYNFMKLKDIAMLKKEIYAIKNDLSRKEKKYIQEI 359

Qy	1081 KILKEKNAELQMTLKLKEESLTKRASOYSGOLKVLAENTMLTSKL--KEKQDKKEILEAEI 1139
Db	360 KSITEINANPEKSVRLNERKMITTVARYSOQLNDLKAENARLSELEKEKHNERLEAEV 419
Qy	1140 ESHHPRLASAVODHDOQIVTSRKSQEPAFHTAGDACLQKKNVDVSVSTIYNNVHLHOPLSE 1199
Db	420 ESHSLSLATANEYEV-ERKOLELVLWRADDSRHEKMGSNISQITDKNELLTEQVHK 478
Qy	1200 AQRKSKLKNLVAGDAIRENTLVSEHAQDQDRETQOCOMKEAHMVQNEODNVNKHTEQ 1259
Db	479 ARVKFNTLKGKIRETRDALREKTLALGVSQDLRQAOHRKEMQKMPNGEAKESQSIGK 538
Qy	1260 QESLDOKLQFQSQNMVLOQLVHAHKADNKSKITIDIH--FLERKMQHLLKQNEEI 1317
Db	539 QNSLEERIQOELNELLERQLEDARKEGDK-EIVINIHRDCLENG-KEDLLERNEKEL 596
Qy	1318 FYNHNLKNRIYQYEKEKAETE 1339
Db	597 MKEYNYLKEKLLQCEKEKAERE 618
RESULT 14	
Q6S8J5 PRELIMINARY; PRT; 712 AA.	
ID	Q6S8J5
AC	Q6S8J5;
DT	05-JUL-2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	POTE2A (Fragment).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	PubMed:15276201; DOI=10.1016/j.gene.2004.05.009;
RT	Bera T.K., Huynh N., Maeda H., Sathyanarayana B.K., Lee B., Paetan I.;
RT	"Five POTE paralogs and their splice variants are expressed in human
RT	prostate and encode proteins of different lengths.";
RL	Gene 337:45-53 (2004).
DR	EMBL; AY42871; AAS58865.1; --
DR	HSSP; P20749; IKLB.
DR	InterPro; IPR002110; ANK.
DR	Pfam; PF00023; Ank; 6.
DR	PRINTS; PR01415; ANKYRIN.
DR	SMART; SM00248; ANK; 6.
DR	PROSITE; PS50088; ANK_REPEAT; 4.
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW	ANK repeat.
FT	NON_TER 712 712
SQ	SEQUENCE 712 AA; 80752 MW; 12777BACA6AF5137 CRC64;
Query Match 8.6%; Score 594; DB 2; Length 712;	
Best Local Similarity 30.4%; Pred. No. 4.2e-14;	
Matches 172; Conservative 106; Mismatches 218; Indels 70; Gaps 19;	
Qy	9 NLNIQDAQKRTALHACVNGHEEVVTFVDRKQQLDVLGDEHRTPLMKALQCHQACANI 68
Db	165 DYNKDKQKRTALHLASANGSEVVKLLDRCOLNLDNKKRTALIKAVQCEDECALM 224
Qy	69 LIDSADINLVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPLLSITKRS 128
Db	225 LLEHGDPNIPDEYGNNTLHYAIYNEDKLMKALLYGADIESKKNHGTPLLLGVHEQK 284
Qy	129 EQIVFELLKNNANAVNKYCTALMLAVCHGSSEIVGMLLQONVDVFAADICGVTAHY 188
Db	285 QQVVKFLIKKANLALDRYGRYGRYGRYGRYGRYGRYGRYGRYGRYGRYGRYGRYGRY 344
Qy	189 AVTCGPHHIEQIMEYIRK--LSKNHNTNPEGSTAGTDEAAPLAERTPDPAESLVEKT 246
Db	345 AVSSHHVVCQLSDYKEKQMLKISSENSNPQDLKLTSEEE---SQRFGKSENSQPEKM 401

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:56:15 ; Search time 37.9263 Seconds
(without alignments)
3402.035 Million cell updates/sec

Title: US-09-602-362E-23
Perfect score: 6920
Sequence: 1 MTRKKTINLIQDAQRKTA.....NHLKNRIYQYEKAETENS 1341

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	431.5	6.2	1738	T14867	interaptin - slime
2	417.5	6.0	2663	1 S28261	centromere protein
3	408	5.9	3259	1 A56539	giantin - human
4	407	5.9	3225	2 I52300	giantin - human
5	406.5	5.9	3488	2 T34418	hypothetical prote
6	403.5	5.8	1875	2 S38173	myosin-like protei
7	400.5	5.8	2116	2 A26655	myosin heavy chain
8	399.5	5.8	2954	2 T14156	kinesin-related pr
9	393.5	5.7	2253	2 T30336	nuclear/mitotic ap
10	387.5	5.6	1313	2 F96673	hypothetical prote
11	385.5	5.6	1957	2 T38077	hypothetical coile
12	381	5.5	3187	2 JC5837	364K Golgi complex
13	377.5	5.5	2139	2 T18296	myosin heavy chain
14	376.5	5.4	5327	2 T13564	microtubule-associ
15	373.5	5.4	2057	2 S61477	myosin II heavy ch
16	368.5	5.3	2017	1 A36014	myosin heavy chain
17	365.5	5.3	1690	2 T13030	microtubule bindin
18	360	5.2	1790	2 S67593	transport protein
19	360	5.2	1964	2 A59282	nonmuscle myosin I
20	358	5.2	1300	2 I53799	CGI protein - huma
21	358	5.2	1410	1 A57013	early endosome ant
22	355	5.1	944	2 S26710	spindle pole body
23	355	5.1	1837	2 T41023	probable nuclear p
24	355	5.1	2168	2 T30171	ninein - mouse
25	353	5.1	2331	2 T25410	hypothetical prote
26	352.5	5.1	1269	2 F84730	probable myosin he
27	350.5	5.1	1935	1 A37102	myosin beta heavy
28	350.5	5.1	2245	2 T18278	myosin heavy chain
29	350	5.1	1356	2 S32763	kinectin 1 - human

30	350	5.1	1938	2 JC5421	smooth muscle myos
31	350	5.1	1972	2 JC5420	smooth muscle myos
32	348.5	5.0	1972	1 A41604	myosin heavy chain
33	348	5.0	1538	2 T29095	cardiac muscle fac
34	347	5.0	1558	2 B71603	RESA-H3 antigen pp
35	347	5.0	1959	1 A33977	myosin heavy chain
36	345	5.0	1940	1 A24922	myosin heavy chain
37	345	5.0	1979	1 S03166	myosin heavy chain
38	344.5	5.0	1961	1 A61231	myosin heavy chain
39	344	5.0	1939	1 A46762	myosin alpha heavy
40	343.5	5.0	1526	2 A45605	mature-parasite-in
41	342.5	4.9	1427	2 S22695	restin - human
42	342.5	4.9	1827	2 T16270	hypothetical prote
43	342.5	4.9	1999	1 S21801	myosin heavy chain
44	341	4.9	1392	2 A43336	microtubule-vesicl
45	340.5	4.9	1940	1 S04090	myosin heavy chain

ALIGNMENTS

RESULT 1

T14867

interaptin - slime mold (Dictyostelium discoideum)

C/Species: Dictyostelium discoideum

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C/Accession: T14867

R/Riveto, F.J.; Kuspa, A.; Brokamp, R.; Matzner, M.; Noegel, A.A.

J. Cell Biol. 142, 735-750, 1998

A/Title: Interaptin, an actin-binding protein of the alpha-actinin superfamily in Dictyo

ts.

A/Reference number: Z18248; MUID:98365468; PMID:9700162

A/Accession: T14867

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1738 <RIV>

A/Cross-references: UNIPROT:O76329; EMBL:AF057019; NID:g3549260; PID:g3549261; PIDN:AAAC

C/Genetics:

A/Gene: abpd

A/Introns: 173/2; 1680/1

Query Match 6.2%; Score 431.5; DB 2; Length 1738;
Best Local Similarity 19.6%; Pred. No. 4.5e-10;
Matches 316; Conservative 270; Mismatches 478; Indels 549; Gaps 70;

Qy 15 AOKETALHWACVNGHBEWTFVLVDRK-----COLDVLDDGEHRTPLMKAL-----OCH 61

Db 23 AOKKVFVWNCNI-----FLNORSQKIEDLETDLVDGILLGSLLEILSGKNVILSKCK 74

Qy 62 Q-----EACANILDSGADINLVDVY-----GNMALHYAVYSEILSVVAKLLS 104

Db 75 QLKTRLHVINNLPFLKFIQDEG--LRLVGVAESDITDGNLKL-----ILGLVWTLIL 125

Qy 105 HGAVIEVHNKASLTPLLLSTKRSQIEVFLLIKVANAVNKYKCTALMLAVCHGSSEI 164

Db 126 RYQIQSQMNSKSSQONLHSTKPS-----LMLNWKVQSIDYG-----HHIKDL 170

Qy 165 VGMLLQONVDFAADICGVTAEHYAVTCGFPHHIEQIMYIRKLSKXNHTNPEGTSAGT 224

Db 171 TTSP--QNGLLFCALV-----HKLVPEKLDY-KSLG-----RSDSLGN 205

Qy 225 PDEAAPLAERTPDPTAESL-----VEKTPDEAAPL--VERTP-----DTAESL-----VEKTP 269

Db 206 LTLAFEVANKELGIPSLDPHDIIITTPDELISILYISLFPKVVQOTLEPLNNNNNISPSL 265

Qy 270 DEASLUVEGTSKIQCLEKATSGKFE-OSAEETPREITSPAKETSEKFTWAKGPRKIA 328

Db 266 SSSSSSLNTPNKNSIQLSKSTSEFQONQQOQNLLSP-----NSYRNSIS 313

Qy 329 WEKKEDTPREIMSPAKETSEKFTWAAKGRPKIAWEKKEKTPVKTGCVARVTSNKTVLEK 388

Db 314 FSK-----SPSFEFSQ-----STGSSRSIS--PISSPIKNSITGNSLNLSKSTSPF- 356


```
Qy 962 VKWQELCSVRLTNQBEK-----RRNADIINEKIRELGRIBOHRKELVKKOOLEAL 1017
Db 1946 TKVESEIRKEYLEIKQAQKPGNKSNAKELQELLKE-----KQBVVKQLQKDCI 1995
Qy 1018 RIQD-----IELKSVEINLQVSHTHENYLLHNCMLKKEIAMKL----- 1060
Db 1996 RYQEKISALERTVVALEFVQTESQ-KDLEITKENLAQAVEHRKKAQAEASFKVLLDDTQ 2054
Qy 1061 -EIA-----TLKHQVQEKENFYFEDIK-LKEKNAELQMTLKLKEESLTKRASVSGQLK 1113
Db 2055 SEAAVLADNLKLKELQSNK--ESVKSQMKQKQEDLERLEQAEEHLKKNQWQEKLD 2112
Qy 1114 VLIENVTMLTSKLE-----KQKEI--LEAEIESHHPLA-----SAVQHDQIVT 1158
Db 2113 ALREKVLHEBTIGEIQTWLNKQKQVQOQENLDSTVTQLAAFTKSMSSLQDDRDRVID 2172
Qy 1159 SRKSOEPAFHITAGDACLQKKNVDVSSITVYNNVHLQPLSEAQKSKSLKINL----- 1212
Db 2173 BAKWERKF-----SDAIQSKEBEIRLEKED--NCSVLKQQLQRMGSHMEELKINIRLEHDK 2227
Qy 1213 --YAGDALRENTL---VSEHAQRDQRETQCCQMKAEHMYQNEQNVNKHTEQOESLQDLK 1267
Db 2228 QIWESKAQTEVQLOQKVCDDTLQGENKELLSQLETRHLHYSSQNELAKLESELKSLKQDL 2287
Qy 1268 POLSKNMWLOOQ-----LVHAHKADNKSKITIDIHFLERKMQ-----HHLKE 1312
Db 2288 TDLNSLEKCKEQKGNLEGIIRQOEADIQNSKFSVE--QLBTDLQASRELTSRLHEEINM 2345
Qy 1313 KNEEIFYNNHKNRI-----YOYEKEKAETEN 1340
Db 2346 KEQKIISLLSGKEBAIQVAIAELRQOQHDKEIKELN 2381

RESULT 5
T34418
hypothetical protein F12F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T34418
R:Fulton, B.; Wohldmann, P.
submitted to the EMBL Data Library, July 1998
A:Description: The sequence of C. elegans cosmid F12F3.
A:Reference number: Z21521
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3488 <FUL>
A:Cross-references: EMBL:U80022; PIDN:AAC25985.1; GSPDB:GN000023; CESP:F12F3.3
A:Experimental source: Strain Bristol N2; clone F12F3
C:Genetics:
A:Gene: CESP:F12F3.3
A:Map position: 5
A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 5.9%; Score 406.5; DB 2; Length 3488;
Best Local Similarity 21.6%; Pred. No. 1.1e-08;
Matches 273; Conservative 217; Mismatches 519; Indels 253; Gaps 48;

Qy 213 QNTNPGTSAGTPDEAAPLAERTPDATABSLVEKTPDEAAPLVER-----TPDTAESL 264
Db 311 QAINVAGEAWCFSDVVMHSESRDDDKSVDE--VDDSTVLEEKKDDGDDKSKPKTKKI 368
Qy 265 VEK--TPD-EAASLVEGSDKIQCLEKATSGKFSQSAEETPREITSPAKSEKFTWPAK 321
Db 369 IKKETPESEQVTAABEPQOKISEVD-----VQSVAEV--EVGAKKPPDAEKPTDLISK 419
Qy 322 GRPKIAWEKKEDTPREIMSPAKETSEKFT---WAAKGRPRKIAWEKKETPVKTCGVARV 378
Db 420 A--KQDSKSKSDP-EASTEESTTEKPTNDKTSKSAEKKTVKPKKVTGVPKLEAKKP 476
Qy 379 TSNKTVLEKRSKMIACPT-----KESSTKA-----SANDQRPSESQOEBEEYSCDSRS 430
Db 477 VEDKDKASQPSSESPPTDGGKKKQIPKFIAPDEITSSRFDPSTHSHSETNITTTIRG 536
```

```
Qy 431 LFESSAKIQVCIPE-----SIYQKVMENRVEEPPKPSAFKPAIE-----MNSVYP 478
Db 537 R-EGSADAKTFLVPLSASVSMKVFTLVESAKE--KAEFSFKRRSETPDDSKRKEGLPP 593
Qy 479 NKAFLKNEQTLR---ADMPFPPEKQKDYENSWDSLSLCTETVSQKDVCLPKATHQKEI 535
Db 594 AKSEKKDEVTAEEKQSTEALI--ESKKEVDSEKISEQQPSDDKNKSEVVGPEKAAGPET 651
Qy 536 DKINGKLEESPNKQGLLKATCGMKVSIPTKALELKMOTPKAEPGPKPSAFEPATEMOKS 595
Db 652 KQDVSEIEEVPKTKIKKTEKSDSSISQSNVLKPADDKSKSDVDVTDKSKKTTEDQTK 711
Qy 596 VPKNALELKNQOTWR---ADEILPSESQKQDYENSWDTSLSL-----CE 636
Db 712 VATDSKLEKAADTTQIETETVDDKSKKVLKKTKEKSDSFISQKSETPPVVEPTKAE 771
Qy 637 TVSQKDVCLPKAAHQKEID---KINGKLEGSFVKDGLLK--ANCGMKVSIPTKALELMDM 691
Db 772 SEAQKIAEVNKAQKQKEVDNLRKAEAEVAAKIADEKLKIEAEANIKKTAEEVAAK-- 827
Qy 692 QTFKAEPEKPSAFEPATEMOKSVPNKALELKNQOTLRADILPSESQKQDYEESSWDSE 751
Db 828 ---KQEKDEBQLEKLETVVSKSAEK--LELEKQAIKKAABADAVKKQELNEKV---- 879
Qy 752 SLCTVSQKDVCLPKATHQKEIDKINGKLEESPNDGFLKAPCRMKVS-----IPT 802
Db 880 -----KLEAAKSAADKL--KLEEE-----SAAKSKVSESVKFGEEKT 918
Qy 803 KALELMDQTPKADPEPKPS-----AFEPATEMOKSVPNKALELKNQOTLRADQMPPS 855
Db 919 KAGE--KTVQVESEPTSKKTTIDTVGATEPADETPKKIIKKKTEKSDSSISQKSDTS 976
Qy 856 E--SKOKKVEENSWDSLSRETVSQKDVCPKATHQKEMD-----KISGKLEBDSLS 906
Db 977 EKVSKQKEQDEPTPAVSETQMTVEAD---KSKKQKETDBKLKLDABIAKTKQEADEK 1032
Qy 907 KILTVHSCEREARELOKHCEQRTGKM-----EQMKKFCVLKKLKSAKEIKSOLE 958
Db 1033 SKLD--AQEKIKKVSDDAARKEKELNDKLKLESEIATKASADKLKLEQAQAKAAE 1089
Qy 959 NQKVWQELCSVRLTNQBEKEKRNADILNEKIREELGRIBEQ-----HRK 1005
Db 1090 VEAARKQKE-----KDEQLKLDTEAASKKAAAEKLELEKQAIKKAAGADAVKKQK 1140
Qy 1006 ELEVKQOLEQALRIQDIELKSVEINLQVSHTHENYLLHNCMLKKEIAMLKLKLEIATL 1065
Db 1141 ELDEKNKLEANKKSAAGKLKIEESAAKSKQTVBEO-----AKLDAQT- 1183
Qy 1066 KHQYQEKENFYFEDIKILKEKNABLOMTLKLKESLTKRA-----SOYSGQLKVLJ-- 1116
Db 1184 KAKTAEQTKLEKDEKSTKESKSETVDEKPKKVLKKTKEKSDSSISQKSETSKTVVES 1243
Qy 1117 -----AENTMLTS---KLKEQKQDKEILEAEIESHHPLRSLASAVQHDQIVTSRKSOEPAPH 1168
Db 1244 AGPSESETQKVADAAARKQKQETDEKQLEAEI-----TAKKSADESK 1285
Qy 1169 IAGDACLQKKNVDVS--STIYNNVHLQPLSEAQKSKSLKINLYAG-----DALR 1219
Db 1286 LEAESKLKKAEEVAAAKKQKEBQLEKLDTEAASKKAAAEKLEKQSHIKKAABEVDVK 1345
Qy 1220 ENTLVSHAQRDQRETQCCQMKAEHMYQNEQNVNKHTEQOESLQDLKFLQSKNMWLOO 1279
Db 1346 KQKEL-BEKQLESEATKKADEKLKLEEQ-----KKKAAEIALIEIQE---QE 1392
Qy 1280 QLVHAHKADNKSKITIDIHFLERKMOHLLKKNBEIFNYYNNHJONRIYQVEKEKATE 1339
Db 1393 KLAQEQSLREDEAKSAEKQKLESETSKQTEEPAKB--SVDEKPKKVLKKTKEKSDSS 1450
Qy 1340 NS 1341
Db 1451 IS 1452
```

RESULT 6

S38173 myosin-like protein MLP1 - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: UNIPROT:Q02455; EMBL:Z28320; NID:G486586; PID:G486587; MIPS:YKR095w
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 03-May-1994 sequence revision 03-May-1994 #text_change 09-Jul-2004
 C:Accession: S38173; S40647; S31207
 R:Baladron, V.; Ballesta, J.P.G.; Bou, G.; del Rey, F.; Esteban, P.F.; Garcia-Cantalejo, submitted to the Protein Sequence Database, March 1994
 A:Reference number: S38158
 A:Accession: S38173
 A:Molecule type: DNA
 A:Residues: 1-1875 <BAL>
 A:Cross-references: UNIPROT:Q02455; EMBL:Z28320; NID:G486586; PID:G486587; MIPS:YKR095w
 A:Experimental source: strain S288C
 R:Bou, G.; Esteban, P.F.; Baladron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; JI
 Yeast 9, 1349-1354, 1993
 A:Title: The complete sequence of a 15 820 bp segment of *Saccharomyces cerevisiae* chromo
 A:Reference number: S40644; MUID:94205265; PMID:8154186
 A:Accession: S40647
 A:Molecule type: DNA
 A:Residues: 1-1875 <BOU>
 A:Cross-references: EMBL:X73541; NID:G450550; PIDN:CAA51948.1; PID:G450554
 A:Experimental source: strain S288C
 R:Kobelling, R.; Nguyen, T.; Chen, E.Y.; Botstein, D.
 Mol. Gen. Genet. 237, 359-369, 1993
 A:Title: A new yeast gene with a myosin-like heptad repeat structure.
 A:Reference number: S31207; MUID:93247549; PMID:8483450
 A:Accession: S31207
 A:Molecule type: DNA
 A:Residues: 1-300, 'A', 302-1875 <KOE>
 A:Cross-references: EMBL:L01992; NID:G171958; PIDN:AAA34783.1; PID:G171959
 C:Genetics:
 A:Gene: SGD:MLP1
 A:Cross-references: SGD:S0001803; MIPS:YKR095w
 A:Map position: 11R

Query Match 5.8%; Score 403.5; DB 2; Length 1875;
 Best Local Similarity 21.2%; Fred. No. 6.7e-09;
 Matches 334; Conservative 258; Mismatches 527; Indels 453; Gaps 73;

QY 40 KCOLDVGEHRTPLMKALQCHOEACANILDSGADINLDVYGNMALHYAVISILSV 99
 DB 124 KKLDLLEE-----KKTQSNQORTKLDERLKEIELVVENRS-----NSEC 169

QY 100 AKLSHGAVIEVHNKASITPLLSITKSEQIVFLLKNAN----- 141
 DB 170 KKLSTIMDLTKQOQYITNDLSNRTELEKTKOELTLQSNNDWLEKELRSKNEOYLSYR 229

QY 142 -----ANAVNKYKTALMLAVCHGSSIVGMLLQONVDVFAADICGVTAEHYAVT 191
 DB 230 QKTDKVIDIRNELRLNRDNFOMERTNN-----DVLKQKNNELSKS----- 270

QY 192 CGFHIIHEQIMEYIRKLSKXNHTNPETGTSAGTDPDEAAPLAERTPTDAESLVEKTPDRAA 251
 DB 271 -----LOEKLE- IKGLS- DSLNSEKQFSA-----ENSLKQRLVDLLESQNLNAVEELN 318

QY 252 PLVETPTDAESLVE-----KTPDEAASLVEGTSKIQCLEKATSKFQSAETPR--EI 305
 DB 319 SI--RELNTAKVIADDSKKQTPENEDLLKE-----LQTKELQAQCEKELRLSSI 367

QY 306 TSPAKETSEKFTWPAKGRPRKTAWEKEDTPEIMSPAKETSEKFTWPAKGRPRKIAWEK 365
 DB 369 TDEADEDNENLS--AKSSSDFIPLKKQLIKERTKEHLNQOITETFIVELEHKVPIINSFK 425

QY 366 KETPV-----KTGCVARTSN-----KTKVLEKGRSKMIACP-----TKES----- 401
 DB 426 ERTDMLNELNNAALLLEHTSNEKNAKVKELNKQKLVCECENDLQTLTKQRLDLCRQIQ 485

QY 402 ----STKASANDORFPSESKQ-----EEDREYSCDSRLFESSAKIOVCIPESIOKV 450
 DB 486 YLLITNSVNSDKGFLRKEIOFIQINIMQEDDS-----TITESDSQKV 528

QY 451 MEINREVEBPCKPSAFKPAIEMQNSVPNKAFELKNEQTLR-----ADPMFPPEKQKOY 505
 DB 529 V-TERLVE-----FKNIIOQE-----KNAELLKVVRNLADKL---ESREKKS 567

QY 506 EENSWSSESCEVTSQ-KDVCL-----PKATHQKEIDKINGKLEE-----SPKQGLLKATC 556
 DB 568 KQSLQKIES--ETVNEAKEAIITLKSEKMDLESRIELOKELBELKTSVFNEDASYNVT 625

QY 557 GMKVSPTKALE--LKDQOTFKAEPKPSAFEPATEMOKSVPNK-----ALE 602
 DB 626 IKQLTETKRDLESQVQLQF-----RISQITRESTENMSLNKEIQDIYDSKDISIK 678

QY 603 LKNEQTRADEILPSE-----SKQKDYEN-----SWTESLCE 637
 DB 679 LGREKSSR---ILAEERFKLLSNTLDLTKAENDQLRKRPDYLTQNTILKQDSKTHETLNEY 735

QY 638 VSQK---DVLCPKAAHQKEIDKINGKLEG-----SPKQGLLKANGCMKVISITKA 685
 DB 736 VSCSKLSIVETELLNLKERQKLRLVHLEKNLKOELNKLSPKDSL-----R 781

QY 686 LELMDQOTFKAEPKPSAFEPATEMOKSVPNKALELKNQOTLRADILPSESQKDY-- 743
 DB 782 IMVTQLQTLQ---KEREDLLE---ETKSCQKKIDELED-----ALSELKETSQKDHII 830

QY 744 ---EESS---W-----DSESLCETVSQKQVCLPKATH-----OKEI--DKIN- 777
 DB 831 KQLEEDNNSNIWQNKIEALKXDYESVITSVDSKQTDIEKLYKVKSLKEKEIEEDKIRL 890

QY 778 ---GKLEESPDNGFLKAPCRMKVSIPT---TKALELMDM-----OTFKASPPKPSAFEP 826
 DB 891 HTYNVMDTINDDSLRKELEKSKINLTDAYSQIKYKDYLTETTSQSLQQTNSKLDESFKD 950

QY 827 AIENQKSVPNKALELKNQOTLRADQMFPPSESK---QKKVEENSWDSLSLETYSQ---KD 880
 DB 951 FTNQIKNLTKETSLEDKISLLKEQMFNLNELLQKGMKEKADFKKISILQNNKE 1010

QY 881 VCVPKATHOKEMDKISKLEBDST-----KRNADIILNEKI----- 992
 DB 1011 VEAVKSEVESKLSKIQNDLQOQTIYANTQNNVBEQLQKHADVSKTISLSEQLHYTKGQ 1070

QY 917 -RARELQKDHCEQRTGMEQMKKFCVLLKKLSAKEL--KSOLEN-----QKVKEQ--- 966
 DB 1071 VKTLNLSRDQLE---NALKENESWSQKESLLEQLDLSNRIEDLSQNKLLYDQIQY 1127

QY 967 -----ELCSVRLTLNOEBE-----KRNADIILNEKI----- 992
 DB 1128 TAADKEVNNSTNGPLANNILITLRRERDILDTKVTVAEERDAKMLRQKISLMDVBLQDART 1187

QY 993 -----REELGRIBEQHRKELEVKQOELQALRIQDIELKSVESNLSNOVSHTHENYL 1044
 DB 1188 KLDNSRVEKENHSHIIQCHDDIMEKLNQL-NLLRESNITLR-----NELENNNNKKEL 1240

QY 1045 LHENCMKKETAMLKLEIATLKHQYQSKENKYFEDIKILKEKNAELQMTLKLKESLTKR 1104
 DB 1241 QSEIDLKQNVAPIESELTKYSNQEBE-----QELKLAKE---EVHWRKRSQDILEKH 1293

QY 1105 ASQYSGQKVLIAENTMLTSLKLEK-ODKEILEAEIESHHPRLASAVQDHQIVTSRKSQ 1163
 DB 1294 EQLSSDYKLESE-----IENLKELENERKQGAEEKFNLRRAQ--ERLKTSLKSQ 1347

QY 1164 EPAPFHAGDACLQKRMNVDSSTIYNNVHLHQPISEAQRKSKSLKINLNYAG-----DAL 1218
 DB 1348 D-----SLTEQVNSLRDAKNVLENSLEANARIEELQ-NAKVAQGNQLEAI 1393

QY 1219 RENTLVSEHAQRDQRETOCQKMEAHMYQNEQDNN-----KHTEQOESLDOKLQFOLQ 1271
 DB 1394 RK---LQEDAKAGRELQAKLEESTSYESTINGLNEEITTLKKEIEKQROIQQOQLOATS 1450

QY 1272 SKNNWMLQOOLVHAHKADNKSKITIDHFLERKQKHLLKEKNEPFIYNNHNL---KNRI 1328
 DB 1451 ANEQNDLSNIVESMKKSFEEDK---IKFIKEKTQ-----EVNEKILEAQERLNQPSNIN 1501

Qy 1329 YOYKKAETEN 1340
Db 1502 MEIIRKKWSEH 1513

RESULT 7
A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N;Contains: myosin ATPase (BC 3.6.4.1)
C;Species: Dictyostelium discoideum
C;Date: 05-Oct-1988 #sequence revision 05-Oct-1988 #text_change 09-Jul-2004
C;Accession: A26655; A24728; S00250
R;Warrick, H.M.; De Lozanne, A.; Leinwand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A;Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A;Reference number: A26655; MUID:87092266; PMID:3540939
A;Accession: A26655
A;Molecule type: DNA
A;Residues: 1-2116 <WAR>
A;Cross-references: UNIPROT:P08799; GB:M14628; GB:M11938; NID:G167834; PIDN:AAA33227.1;
R;DeLozanne, A.; Lewis, M.; Spudich, J.A.; Leinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A;Reference number: A24728; MUID:86016788; PMID:3901008
A;Accession: A24728
A;Molecule type: mRNA
A;Residues: 2035-2116
R;Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A;Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium myosin heavy chain
A;Reference number: S00250; MUID:88112226; PMID:2828113
A;Accession: S00250
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1734-1893 <WAG>
C;Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphate binding site
F;1-818/Domain: globular head <HED>
F;89-747/Domain: myosin motor domain homology <MMOT>
F;179-186/Region: nucleotide-binding motif A (P-loop)
F;819-2116/Domain: alpha-helical rod <ROD>

Query Match 5.8%; Score 400.5; DB 2; Length 2116;
Best Local Similarity 20.2%; Pred. No. 1e-08;
Matches 266; Conservative 253; Mismatches 470; Indels 329; Gaps 55;

Qy 226 DEAPLAERTPTAESLVEKTPDEANPLVE--RTPTAESLVEKTPDEANSLV---EGTS 280
Db 908 DEKKLALENQKRSVEEKVRDLEELQEEQKLRNTLEKLUKKYBELEEMKRVNDGQS 967

Qy 281 DKIQLEKATSGKPEQSAEETPRITSPAKETSEKFTWPAKG-----RPRKIAWEKKEDT 335
Db 968 DTISREKID-----ELQKEVEBELTESFEESKDKGVLEKTRVRLQSELDL 1015

Qy 336 PRETMSPAKETSEKPTWAAGRPRIKIAWEKK-----ETPVKTCGVARTSNK----- 382
Db 1016 TVRLDSETKKSELRL-----QKKLEELKQVQEAALAAATAAK---LAQEAANKLQGE 1067

Qy 383 -TKVLEGRSKWIACTPKESSTK-----ASANDORPPSESKQBED----- 421
Db 1068 YTELNEKFNSEVTARSNNVEKSKTLESLQVAVNNEL--DEBKNRDALEKKKKKALDAMLE 1125

Qy 422 -----EYSCDSRLSPFASSAKIQVCIPEIYQKVMENREVEBPFPKPSAFKPAIEMQN 475
Db 1126 EMKQLESTGGEKSLVDLVKVK-QESDMEALRNQISELQSTIAKLEIKST-----LEG 1178

Qy 476 SVPNKAFELKNEQTLRADPMPFPSPKQKDYENSWDSLSLCTVTSQKDVCLPKATHOKEI 535
Db 1179 EVARLQGELEAEALQAKSNVKKQKKKVDLELDK---SAQLAEETAAK-----QAL 1225

Qy 536 DKINGKLEESPNKGLLKATCGMKVSPITKALELKMOTFKAEPPGKPSAFEPATE-MQK 594
Db 1226 DKLKKKLEQ-----ELSEVQTLSEANNKNNVNSDSTNKHLET 1262

Qy 595 SVPNKALELNEQOTWR-----ADEILPSESQKQD-YEENSMDTESLCEI 637
Db 1263 SFNNLKLELEAEQAKQKALEKRLGLESELKHVNEQLEEEKKQKESNEKRVLDLK--EV 1320

Qy 638 VSQKDVCLPKAAHOKETDKINGKLEGGSPVQGLKANGCMKVSIPTKALELMDMOTFKAE 697
Db 1321 SELKQDIEEVEASKAVTEAKNKE-SELDS--IKROYADVSSRDKSVE--QLATLOAK 1375

Qy 698 PPEKPSAFEP-----TEMOKSVNKALELNEQTLRADEILPSESQKQD 742
Db 1376 NEELRNATAEAGQGLDRAERSKKAEPDLEBAVN--LEETAKKVKAEM--KKAETD 1431

Qy 743 YEESWDSLSLCTVTSQKDVCLPKATHOKETDKINGKLEGGSPDN-DGFLKAPCRMKVSIPT 801
Db 1432 YRSTKSELDDAKNVSSQYVQIKEL--NEELSELRSVLEEDERCNSAIKAKKTAESALE 1489

Qy 802 TKALELMDMOTFKAEPPKPSAFPAI-EMOKSVNKALELNE-----QTLRADQ 851
Db 1490 SLKDEIDDAANNAKAKAEKSKLEVRVAELSESLEDSKGTYNVVFIRKDAEIODLRAL 1549

Qy 852 MFPSESQKQKVEENSWDSLSLRETYSQKDVCPKATHQKE--MDKISGKLE-DSTSLSK 907
Db 1550 DRETESRIKSD-----DKNTRKQFADLEAKVEEA--QREVVTIDRLKKLESIIIDUST 1603

Qy 908 ILDT-----VHSCERARELQKHCEQRTGM-----EQMKKRF 941
Db 1604 QLDTEKTSRIKIEKSKKLEQTLAERRAAEBGSKAADEBIRKQVQWQVDSLRALDSE 1663

Qy 942 -----VLKKLSEAKETKSQLENO-----KVKWEQLCSVRLTLNQBEEKRR 983
Db 1664 AALNASEKKIKSLVAEDEVQEQLDEILAKDKLVKAKRALEVELEVRDQLEBEDSRS 1723

Qy 984 NADILNEKIREELGRIBEOHRKELEVKQOLEOALRIQDIELKSVESNLNOVSHTHENNY 1043
Db 1724 ELEDKRLITTEVEDIKKYDAEVEQNTKLEA-----KKLTDDVDTLKKQLEDEKK 1776

Qy 1044 LLHENCMLEKEIA-----MLKEIATLKHQYQEKENKYPF-DIKILKEKNAELQMTLKL 1096
Db 1777 KLNESERAKKLESENEDFLAKDAEVKNRSRAEKDRKKYKEDLKDTK-----YKL 1827

Qy 1097 KEESLTKRASQY-SGQLKVLIAENTMLTTSKKEVQ-----DKLEIAEIESHHPLA 1147
Db 1828 NDEAATKTQTEIGAALKEDQIDE---LRSLKEQQAQATQADSKKTKLEGGIDN---LR 1880

Qy 1148 SAVQDHQIIVTSRKSQSPAFHAGDACIQRCQNVDSSTIYNNVHLQPLSEAOQKSKSL 1207
Db 1881 AQIEDEGKIMRLSEKRA-----LEGELE-ELRETV---EAEDESKSEAEQSKSLV 1928

Qy 1208 KINL-----NYAGDA-----LRENTLVSEHAQRDQRETQCOQ 1239
Db 1929 ELELEDARRNLQKEIDAKEIAEDAKSNLQREIVEAKGRLEBESIARTNSDRSKRLEAEI 1988

Qy 1240 K-----EAEHYQNEQDNVNHKTEQOESLDQKLF-----QLQSKMMVLOQLVHAHKA- 1288
Db 1989 DALTAQVDAQAKNQOIKENKKTIELKEYRKFGSESEKTKTKEFLVVEKLETDYKAK 2048

Qy 1289 -----DNKSKITDIHFLERKQMHLLKEKNEEIFNYYNNHLKNRIYQVEKEKAETENS 1341
Db 2049 KEAADEQQORLTVENDLRKHLSEISLLKDAIDKL-----QRDHDKTKKELETETA 2098

RESULT 8

T14156

kinesin-related protein - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C;Accession: T14156

R;Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.

Cell 91, 357-366, 1997

A;Title: CNP-B is a plus end-directed kinetochore motor required for metaphase chromosome

A;Reference number: 217893; MUID:98028574; PMID:9363944

A;Accession: T14156

A;Status: preliminary; translated from GB/EMBL/DDB3
A;Molecule type: mRNA
A;Residues: 1-2954 <WOO>
A;Cross-references: UNIPROT:O42263; EMBL:AF027728; NID:g2586070; PID:g2586071; PIDN:AAAC
C;Genetics:
A;Gene: XCENP-B
C;Superfamily: centromere protein E; kinesin motor domain homology

Query Match	5.8%;	Score 399.5;	DB 2;	Length 2954;
Best Local Similarity	20.0%;	Pred. No. 1.7e-08;		
Matches	304;	Conservative 269;	Mismatches 575;	Indels 369; Gaps 62;

QY 11 NIQDAQKRTALHWACVNGHEVVTFFVDRKCOLDLVDGHEHRTPLMKALQCHQACANILI 70
DDB 1195 DLEGEVKKLLLEMLLKGH-----LTDQSLSIEKLEQLN-LEVTEKLTQLQEMKNYITI 1247

QY 71 DSGA-----DINLDVGVGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLT 118
DDB 1248 ERNELQTNFEDLKAHDSLSLQDLSENIEQSIQTQDELRAAQAEELREQQLVDSF-RQOLL 1306

QY 119 PLLSITKRSQIVEFLIKNANAVNKYKCTALMLAVCHGSSEIVG---MLLQNVDV 175
DDB 1307 DCSVGISSPHNDAY-----ANQEKVSLGEVNSLOSEMLRGERDELQTSCKALVSELEL 1359

QY 176 FAADICGVTAHYAVTCGFHHIHQIM-----EYIRKLSKNHONTNP-----217
DDB 1360 LRAHVSVGEENLEITKLNGLKLEIKLGSSESVLKSMLNKEDNNKLKEQAEEYSSK 1419

QY 218 -----EGTSAGTP---DEAAPLAERTPTDAESLVEKTPD-----EAPLVE---RTP 258
DDB 1420 ENQFSLBEVSGSKLVDIEVLKAQLKAAAEERLEIKDRDYFELVQVQANTNLVEGKLETP 1479

QY 259 ---DTASLVEKTPDEAASLVEGSDKIO---CLEKATSGKPEQSAEETPREITSPAKE 311
DDB 1480 LQAHEDSDIRREEMIEKVLG---EKLENOYLLERLQEKLELS---NKLELQKEME 1534

QY 312 TSEKFTWPAKGRPRKIAWEKEDPTPEIMSPAKTSBKFTWAAKGRPRKIAWEKFTPVK 371
DDB 1535 TSVLL-----KDDLQOKLESLLSENI-----ILKENIDITLTK 1566

QY 372 TGCVARVTSNKT-KVLEKGRSKMTA---CP-TKESSTKASANDORPPSESK---QEED 421
DDB 1567 HHSPTQAOQLKQTQOELQAKNLAAASNCPIQEKETSA---DCVHPEEKILLILLTEEL 1623

QY 422 EEYSCDSRLSPSSAKIOVCIPESYIQKWMENREVEEPPKPSAFPAEMQNSVPNKA 481
DDB 1624 HQTNEQEKLLHKEKNELE-----QAQVELKCEVEH-----LMKSMIESKSLSLQ 1669

QY 482 FELKN-EGTLRA-----DMFPPEPK--QKVYENSWDSESLCETV-----SOKDVCLPKAT 530
DDB 1670 HEKHDTQQLLALQKQOMVVTQEKKELOOTHEHLTAEVDHLKENIELGLNFKNEAQOKTT 1729

QY 531 HQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELKDQMTFKAEPPKPSAFAPAT 590
DDB 1730 KEQCLNLENKELEQSHR-----LQCEIEELMKSLKD-----KESALETLK 1770

QY 591 EMOKSVPN-----KALELKNQETWRADEILPSESKQKVYENSWDSESLCETVSQ 640
DDB 1771 ESEQKVINLQEMEMVWMEELKNSQ---RTVIAERDQLQDDLRES---VEMSIET--Q 1822

QY 641 KDVCPLPKAAHQKEIDKIN-----GKLEGSFV-----KDGILLKANC 675
DDB 1823 DDLRKAQEAQOQKQKQVOLTQSIVLQEKISLLENQMLYVAVTKETLSERDDLNSQSQ 1882

QY 676 GMKVSIPTKALELMDQMTFKAEPPPE--KPSAFPAIEMQKSVPNKALELKNQETLRADEI 733
DDB 1883 HLFSFIEITLSLSKEKE-FALQAEKDKADAARKTIDITEKISNTIEEQQLQOATNLKETL 1941

QY 734 LPSSESKQKDYESSWSDSESLCETVSQKDVCLPKATHQKE-----IDKINGKLEESP 784
DDB 1942 YRESLQCKEQLALNTEHLRETLKSLDALQKMEQERDEAANKVIALTEKMSLSEEQIN 2001

QY 785 DNDGFIKAPCRMKVSIPTKALELMDQMTFKAEPPPEKPSAFPAIEMQKSVPNKALELKN 844

Db 473 YEKMLEDARNEIDSLKS-----TVDSIQNEFENS KAGWEQKELHLMGCVKKEE-----NS 524

QY 530 THOKIDKINGKLESPNKGDLK--ATCGMKVSIPTKALELKMOTKFAPPGKPSAF- 586

Db 525 SSQEEVSRKLVNLLKES-BEDACARKEEASLKNLKVAGGEYKYLQETIGAKAESMKL 583

QY 587 EPATWOKVFNKALELNEQOTWRADEILPSPSKODYEENSWDTESLCTVSKQDVCLP 646

Db 584 ESLDKEDLKNVTAEISLREWEGSVLEKIEELSK-VKESLVDKETKLQISITQABEELK 642

QY 647 --KAARHOKIEIDKI-----EGSPVKD-----GLLKANGCMKVSIPTKAL 686

Db 643 GREAAHMQIEBELSTANASLVDATKLIQSVQESDLKEKEAGYLLKTEELSVANESLAD 702

QY 687 ELMDMOTKFAEP---PEKPSAFPAEMOKSVFNKALELKNQOTLR-----ADEILPSE- 737

Db 703 NYTDLQSVQESKOLKEREVAYLKKIE-ELSVANESL-VDKETKLQIHDOAEELRGREA 760

QY 738 SKQKDYESSWDSLSCTVS-----QKDVCLPKATHOKEIDKI---NGKLEESPDN 786

Db 761 SHLKIEBELKENENLVNANVMQNIABESKDLREREVAYLKKIDELSTANGTLDADVTN 820

QY 787 DGLKAPCRMKVSIPTKALELMDMOTKFAPEKPSAFPAEMOKSVFNKALEL----- 841

Db 821 -----LQNISEENKELRERETLLKKAELS-----ELNESLVDKASKLQTVVQ 864

QY 842 KNEOTLRADOMPSPSKOKKYEENSWDSESLBETVSQKDVCPKATHOKE--MDKISGL 899

Db 865 ENEELRERETAY-----LKKIEELS-----KLHEILSDQETKLQISNHEKEBELKERETAYL 915

QY 900 EDSTLSKILDTVHSCERARELQKDHCEORTCKMEQMKKFCVLKKLSEAKEIKSQL-- 957

Db 916 KKBELS KVOEDLLN-----KENELHGMVVEIEDLRKSKSLAOKKIEELSNFNASLLI 968

QY 958 -----ENQKWKVQ-----ELCSVRLTLNQEFEEKRRNADILNEKIRBE----- 995

Db 969 KENELQAVVCENEELSKQVSTLKTIDELSDLKQSLIHKEKELQAAI VENEKLEKAAALS 1028

QY 996 LGRIEFQRKLEVKVQKLEQALRIODIELKSVESNLQVSHTHENENYLLHNCMLKKEI 1055

Db 1029 LQRIEE-----LTNLKQTLIDKQNELQV-----PHNEELKAKEBASLKKID 1071

QY 1056 AMLKLEIATLKHQYQENKYPEDIKILKEKNAELQMTLKEESLITKRASOYSQLKVL 1115

Db 1072 ELLHLE-----QSWLEKSEF-----QRTQENLELKTQDALLAKKIEELSKLES 1118

QY 1116 TAENMTLSKLEKQKQKLEIAEIESSHPRLASAVQDHDQIVTSRKSOEPAPHIAGACL 1175

Db 1119 LEKETEL--KCREAAALEKMEBPSKHGNSLSICKDYDLVQFS-----EVNGASNG 1168

QY 1176 QKQMVVDVSSITYNNEVLHPLSEARQSKSLKINLNTAGDALRENTLVSEHAQDQRET 1235

Db 1169 DEKTKTDHYQORSRHEMTIQESPEAIDK-----HLMGRAAITHKVAHREGER 1216

QY 1236 QCMQKEAE-----HMYQNEODNVNKHTEQ-----QBSLQKLPOLQSKNNMLQQQLVHAHK 1286

Db 1217 NVE-KESFEKMWDSYKIEKSEVSPRETELDSVEBEVDSK--AESSENW---DQYNSGFS 1270

QY 1287 KADNKSKITIDHFLERKQMHLLKKEKEEIEFNPNYNNHLLKNRIYQYEK 1335

Db 1271 LTUDH-----TEDSGNLLLEQH--MKKKPLLRKFGNLLKKKSTSSSQK 1313

RESULT 11

T38077
hypoetical coiled-coil protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38077
R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z21767
A:Accession: T38077

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1957 <CON>

A:Cross-references: UNIPROT:Q10411; EMBL:Z70690; PIDN:CAA94624.1; GSPDB:GN00066; SPDB:SP

A:Experimental source: strain 972h-; cosmid c1f3

C:Genetics:

A:Gene: SPDB:SPACIF3.06C

A:Map position: 1

Query Match 5.6%; Score 385.5; DB 2; Length 1957;

Best Local Similarity 19.4%; Pred. No. 3.8e-08;

Matches 292; Conservative 270; Mismatches 574; Indels 369; Gaps 61;

QY 71 DSGADINLVYGNMHLHYVYSEILSVVAKLLSHGAVIEVHNKASLTPLLLSITKRESEQ 130

Db 62 DSSSELKQKVRG-MRRH-----SDLSIDAKLGS-----SEGSTASSALPLTPRSPS 107

QY 131 IVEFLLIKNA-----NANAVNKYKCTALMLAVCHGSSEIVGMLLQON--VD 174

Db 108 NASMLLVGGLLDSPILDINSVTQKSNLLNELKQVRSKLAALHEHNGILSLQSSSKKD 167

QY 175 VPAADICGVTAHVAVTCGFHHIHOIMYETRKLSKNHNTNPGETSAGTDPDEAPLAER 234

Db 168 KNTSSVTLTSE-----EDVSYFQKLL-----TNMESNFAKQSEAYDLGRQ 209

QY 235 TPDTAESLVEKTPDEAAPLVERTPDTAESLVEKTPDEAAS--LVEGTSDKIQCLEKATSG 292

Db 210 LLTVTEKLDKKEKD-----YEKIKEDVSSIKASLAEOQSNKSLRGEORLEKL-LVSN 263

QY 293 KPEQSAEETPREITSPAKETSEKTPWA-KGRPKRIAWKEK-----ED 334

Db 264 KTVSTLRQTENSLRAECKTLQEKLEKCAINEEDSKLEELKHNVANYSDAIVHKDKLIED 323

QY 335 TPREIMSPAKETSEKFTWAAKGRPKRIAWKEK-----TPVKTGCVAR 377

Db 324 LSTRISEFDNLKSERDTLSIKN-----EKLEKLLRNTIGSLKDSRTSNQSEEMVEL 376

QY 378 VTSNKT--KVLKGRSKMIACPTKESSTKASANDQRPSPSEKQBEDEYSQDSRSLFESS 435

Db 377 KESNRTHTSQTDAESKLSSEFEQENKSLKSIDEVQNNLSKDKMVKVQVSSQLEARS 436

QY 436 AKIQVICPESYQKWEINREVBPPPKPSAF-KPAIEMQNSVPNKAFELKNQOTLRADP 494

Db 437 AHAT-----GKLAEINSEKDFQNKIKDFEKIEQDLQACLNSSNELKESAL---- 484

QY 495 MPPESKQDYENSWD-----SESLCTVSW--QKDVCLPKATH--QKEIDKING 540

Db 485 -----IDKKQQLANLRQIKQEKVSESTOSSQSLQORDILNEKKKHVEVYESQLNELKG 539

QY 541 KL-----EESPNKDGILLKATCGMKVSIPTKALELKD 571

Db 540 ELQTEISNSEHLSQSLTLAAEKEAAVATNNELSEKSNLSQTLQNAFQEKLAKSVMLKE 599

QY 572 -MOTPKABPPGKPSAFEPATEMOKSVPN-----KALELKNQOTWRADEI 614

Db 600 NEQNFSS-----LDTSFKKLNEHQELENHNTITTKQLKDTSSKLQQLQLEAFNFEQKEST 655

QY 615 LPSES-----KQDYENSWDTESLCTVSKQDVCLP-----KAAHQ 651

Db 656 LSDENNDRTLKLLKEESNKSLLIKQEDVDLSLEKNIQTLKEDLRKSEEARPSKLEAKNL 715

QY 652 KE-IDKINGKLE-----GSPVKDGLLKANGCMKVSIPTKALE-----LMDMQTFKA 696

Db 716 REVIDNLKGTLETAQNRDLHSSLSDA--KNTNAILSELSTKSSDEKVLTAANVTLTQ 773

QY 697 EPPEKPSAFPAIEMOKSVFNKALELKNQOT---LRADEIILPSESQKQDYESSWDSSEL 753

Db 774 DSKAMQFSTSLVNSYQSIISNLXHELDRDHVMQSNNTLLESESKL-----TD 823

QY 754 CETVSKQDVCL-----PKATHQ-----KEIDKINGKLEESPDNGFLKAPCRMKVS-- 799

Db 824 CENLTQQNMTLIDNVQKLMHMHVQESKVSSELKEVNGKLSLDLKN---LRSLNVAISDN 880

Db	1461	ESQVSVQNEQDALLGLKALLQERDKLIIVEMDKSLLENQSLGSCSLSLKALUGGLTEDK	1520
Qy	903	TSLSKILDTVHSCER---ARELOKDHCBQRTGRMEQMKKFCVLKXKL-----SEAKEIKS	955
Db	1521	EKLKELSVR-CSKIAESTWQEKH-----KELQKEYEVLLQSVNVSNEARIQH	1571
Qy	956	QLENQKWEQELCSVRITLNOEKEKRNADILNEKIREELGRIBEOHRKELEVYKQOLEQ	1015
Db	1572	VVESVR---QEKQEVYAKLASAESDKRERKQLQDAQEEMEEMKMKRPAKSKQ--K	1625
Qy	1016	ALRIODIELKSVESNLNOVSHENYLLHENCMLKKEIAMLKLEIATLKHQYOEXENK	1075
Db	1626	ILEEE-ENDRLABAQVPGGANSMEALLSNASLKEELRITILEYKLSKEFE-----	1679
Qy	1076	YPEDIKILKEKNAELQMTLKLK---EESLTKRAS-----QYSGOLKVLIAENTMLTSKLKE	1128
Db	1680	-----ALMAEKVTLSEETRNKLQVEAQELKQASLETTEKSDPKDVIIEVTEAVVGKSQ	1734
Qy	1129	KQDKELAEIB-----SHHPRIASAVODHDQIVTSRKSQBPAPHIAGDACLQKRM	1179
Db	1735	EQDSLSENAKLEDAEATLANSAPGVSETTFSSHDDINNLYLQDDQLKGRIAELEMEKQK	1794
Qy	1180	NVDVSTIYNNEVLHQLPSEAKSKSLKI-----NLNYAGDALREN-----TLVSEHA-	1228
Db	1795	DRELSQTLNEK--NALLTQISAKSELKLEEEVAKINMLNQIOEELSRVTKLETAEE	1852
Qy	1229	ORDQRETQCMKEAE-----HMYQNEQDNVNHTEQOESLDQKLOLOKSNMMLQOOL	1281
Db	1853	BEKDDLEERLMQALNELSGNIGYQDVTDAQIKN-EQLESEMQLKRCVSELEBEKQOL	1911
Qy	1282	VNAHKADNKSKITDIHFLER-----KMOHLLKKEKNEI	1317
Db	1912	VKEKTQVSE---IRKEYMEKIQGAQPGSGKIHAKELQELLKKEKQEV	1957
RESULT 13			
T18296			
myosin heavy chain - Entamoeba histolytica			
C;Species: Entamoeba histolytica			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C;Accession: T18296			
R;Guillen, N.			
submitted to the EMBL Data Library, February 1997			
A;Reference number: Z18865			
A;Accession: T18296			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 1-2139 <GUI>			
A;Cross-references: UNIPROT:Q07569; EMBL:L03534; NID:g1850912; PID:g1850913; PIDN:AAB480			
C;Genetics:			
C;Gene: mhca			
C;Superfamily: myosin heavy chain; myosin motor domain homology			
F;91-780/Domain: myosin motor domain homology <MMO>			
Query Match 5.5%; Score 377.5; DB 2; Length 2139;			
Best Local Similarity 22.0%; Pred. No. 8.9e-08;			
Matches 278; Conservative 179; Mismatches 469; Indels 335; Gaps 52;			
Qy	207	KLKSNHONTPEGTSAGTPDEAPLAERTPDPTAESLIVETKTPDPAALPVRTPTAESL--	264
Db	967	EITELNSQINTLNATVNDKDTIAEMQESIDEXEDEITKLKGIKLEEEKDDLEQDRAD	1026
Qy	265	VEKTPDEAASLVEGTSKIOCLE-KATSGKFEQSAETP---REITSPAKETSEKTPWA	320
Db	1027	VSATKDDIAKLNKIT--ICEBAKDIKLEQLEDEENKNDLNELQOOLKGETE	1084
Qy	321	KGRPRKIAWEKEDTPEIMSPAKETSEKFTKTAAGKRPRIAWKKEKTPVKTCVARVTS	380
Db	1085	KSLLAAQVAATKASDERDITLSONLE-NEKLT-----TK	1116
Qy	381	NKTKV---LEKGRKNIACTPKESSTKASANDQRPSPESKQERDEEYSCDSRLFPSSAK	437

RESULT 14

microtubule-associated protein homolog - fruit fly (Drosophila melanogaster)
 N:Alternate names: hypothetical protein EG:49E4.1
 C:Species: Drosophila melanogaster
 C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
 C:Accession: T13564
 R:Spanos, L.; Papagiannakis, G.; Siden-Kiamos, I.; Louis, C.
 A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
 A:Reference number: Z17689
 C:Accession: T13564
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-5327 <SPA>
 A:Cross-references: UNIPROT:O76891; EMBL:AL031128; PIDN:CAA20006.1
 C:Genetics:
 A:Cross-references: FlyBase:FBgn005392
 A:Introns: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1
 A:Note: EG:49E4.1
 C:Superfamily: Drosophila 576K microtubule-associated protein homolog

Query Match 5.4%; Score 376.5; DB 2; Length 5327;
 Best Local Similarity 19.8%; Pred. No. 2.9e-07;
 Matches 279; Conservative 235; Mismatches 595; Indels 297; Gaps 58;

114 KASLTPLLSITKRSQIIVFLLIKNANAVKYKCTALMLAV---CHGSS-----I 164
 2772 RESKPLDSKTSRPGSVESVTADEKSEQSRRESVAESVKADTKDGKSGQASRRPS 2831
 165 VGMLLQNVVFAADICGVTAHYAV--TCGFHHIHEQIMEYIRKLSKHNQNTNPGTS-- 221
 2832 VDELLKDDDEKQESRRQSIITGSHKAMTWG---DESPMD---KADKSEKSPRESVAES 2884
 222 ---AGTPDEAAPLAERTPDTAESL-VEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVE 277
 2885 IKHENTKDEESPLGSRSDVAESIKSDITKGEKSPKSPKSVSRPESVGVGSKDEKA---E 2941
 278 GTSKIOCLEKATSKPQSA-----EETPREITSPAKETSEKFTWPAKGRPKTIW--- 329
 2942 SRRESVAESVPESKDATSPPSKEHSPRESVLSGLKDEGKIT-----SRVSVADS 2995
 330 -----EKEDTPREIMSPAKETSEKFTWAAKGRPKTIWAKETTPVKTG 373
 2996 IKDEKSLVQASRSEASESLKDAAPSOETS-----RPESVTSVKDG--KSP 3044
 374 CVARVTSNKTQVLEKGRSKMTACTPKE-----SSTKASANDORFPSESQDEDEYSC 426
 3045 VASKEASRPASVAENAKDS--ADESKEQRPESLPQSKAGSIKDEKSPKSLASKEAEKSK-- 3100
 427 DRSILFESSAKIQVCIPESIYQKWEINREVEEPKPSAKPALEMQNSVPNKAPELKN 486
 3101 -----EESRR-----ESVAEQPLVSKESVSRP-----ASVAESVKDEAEKSK 3138
 487 BOTLRADPMFPPESS---KQKDYENSMDSESLCETVSO-----KDVCLPKATHQKEDIX 537
 3139 ESPLMSKEASRPASVAGSVKDEAEKS--KEESRRRESVAESKSPKSPASVAESVKDE 3197
 538 INGLKSPNKDGLLKATCGMK-VSIPPKALE-LKDMOTFAEPGPKPSAFEPATEMQKS 595
 3198 ADKSEESRRRESGAESKSPASKEASRPASVAESIKD-----EAEKSEKESRRRESVAE-KSP 3252
 596 VPKALELKNQOTWRADEILPS--ESKQDYENSMDSESLCETVSOQKDVCLP----- 646
 3253 LPFSK-----EASRPTSVAKSVKDEAEKSEESRRSDSVKSPKSPASVAESV 3305
 647 --KAAHQKEIDKINGKLEGSVPKDGLLKANCGMKVSIPTKALELMDMOTFAEPPEKPSA 704
 3306 QDEAEKSEESRRRESVAESKSP-----NYKEASRPASVAESIKDEAEKSEESR--- 3354
 705 FEPATEMQKSPNKALELKNQOTLRADILPS--ESKQDYENSMDSESLCETVSOQKDV 762

3355 -----RESVAEKS-PLASKEASRPTSVAESVKDEAEKSEESRRSDSVKASPLASKEA 3406
 763 CLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMK-VSIPTKALELM--DMOTFFKAEP-P 818
 3407 SRPASVAESVQDEAEKSEESRRRESVAESKSPKSPASKEASRPASVASKDDAEKSEESRR 3466
 819 EKPSAFEPATEMQKSPNKALELKNQOTLRADPMFPSESQKQKVEENS-WDSSESLRETYS 877
 3467 ESVAEKSPLASKEASRPASVAESVKDEAEKSK-----ESRRRESVAESKSPKSPASRPTS 3522
 878 QKDVCPKATHQKEMDKISGLKEDSTSLKILDTVHSCERAR-----ELQKHCE--QRT 930
 3523 VAESVKDEAEKSEESRRRESVAESKSLASK-----EASRPASVAESVKDEAEKSEKE 3574
 931 GKMEQMKKKFVLLKKLSE---AKEIKSOLENQKVKWEQELCSVRLTLNQEEERKNAD 986
 3575 SRRESVAEKSPLASKEASRPASVAESVKDEAEKSEKESRRRESVAESKSPKSPASRPTS- 3633
 987 ILNEKIREELGRIBEOHRKE-----LEVKKOOLEQALRIQIDIELKSVESNLNQVSH 1036
 3634 -VAESVKDEADKSEESRRRESGAESKSPASMEASRPTSVAESVKDETEKSE----- 3684
 1037 THENENYLLHNCMLKKEIAMLKLEIA-TLKHYQE-KENKYPEDIK-----ILKKNAB 1089
 3685 --ESRRSVTEKSPKSPASRPTSVAESVKDEAEKSEESRRRESVAESKSPASKESSRP 3742
 1090 LQMTLLKKESSJ-TKRASQ-----YSGQLKVLIAENTMLTSLKKEQKQKILEAEIESHH 1143
 3743 ASVAESIKDEAGTKQESRRSEMPESGNAESIKGQSSLASKETSRPVSVESVKDETEK 3802
 1144 PRLASAVODH-----DOIVTSRKSOEPAFIAIGDACLQRKNV--DVSSTIYNNVHLQ 1195
 3803 PE-GSAIDKSVASRPESVAVSAKDEKSPKSPASRPSVADKSPDASKEASRSLSVAETASS 3861
 1196 PLSSAQRKSKSLKINLNYAGDA-----LRENTLVSEHAQRDOR----- 1233
 3862 PIEBGRSIADLSPLNLTGEAKGLPTLSPPIDVAEGDFLEVKAESSPRPAVLSPKPAEF 3921
 1234 -----ETQCKMEAEHMYON-----BQDNVNGHT 1257
 3922 SQPDGTHTASTPVDASPVLEIEVVEQHTTSGVGATGATATDLDLDTETKSETVTVKOS 3981
 1258 EQ-----QESLDQKLFQLOSKNMWLOOLVHAHKKADNKGSKITIDIHFLERKMOHLIK 1311
 3982 ETTLFETILTSVESKEVLESSVKQEBKQTSVKQAB--TTVTDLSQLTKSKSEQLTE 4039
 1312 EK---NEEIFNYYNHLKNRIYQYEKE 1334
 4040 IKSVDLTNISVNTLNFSTAVETIEKK 4065

RESULT 15

S61477
 myosin II heavy chain, non-muscle - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
 C:Accession: S61477; S65349
 R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.
 J. Mol. Biol. 255, 98-109, 1996
 A:Title: Molecular organization and alternative splicing in zipper, the gene that encodes
 A:Reference number: S61477; PMID:96144835; PMID:8568878
 A:Accession: S61477
 A:Molecule type: DNA
 A:Residues: 1-2057 <MAN>
 A:Cross-references: UNIPROT:Q94987; EMBL:U35816
 R:Mansfield, S.G.; Al-Shirawi, D.Y.; Ketchum, A.S.; Newbern, E.C.; Kiehart, D.P.
 submitted to the EMBL Data Library, September 1995
 A:Reference number: S65349
 A:Accession: S65349
 A:Molecule type: DNA
 A:Residues: 1-1908, 'NL', 1911-2057 <MAW>
 A:Cross-references: EMBL:U35816; NID:g1141789; PIDN:AAB09049.1; PID:g1572481

Genetics:	
Gene: zip	
Cross-references: FlyBase:FBgn0005634	
Introns: 10/3; 90/3; 162/3; 215/1; 255/3; 263/3; 303/3; 334/3; 1248/1; 1349/3; 1526/3;	
Superfamily: myosin heavy chain; myosin motor domain homology	
Keywords: alternative splicing; ATP; nucleotide binding; P-loop	
135-855/Domain: myosin motor domain homology <MMOT>	
225-232/Region: nucleotide-binding motif A (P-loop)	
Query Match	5.4%; Score 373.5; DB 2; Length 2057;
Best Local Similarity	20.2%; Pred. No. 1.2e-07;
Matches 252;	Conservative 239; Mismatches 427; Indels 331; Gaps 53;
Qy	230 PLAERTPTDAESLVKTPDEAAPLIVERTPDTA-----ESLVKTPDEAASIVEG 278
Db	927 PLLEVYTKQ-EELKQVK-EDELQVREKLTDLAKNTQYERYKYYQALVEKT-----TLAEQ 979
Qy	279 TSDKIQCLEKATSGKF-----EQSAETPREITSPAKETSEKFTWPAKGRPKIAWEKKE 333
Db	980 LQAEIELCAEAESRSLMARKQELDMQOELETRIEEEEE-----RVLAGGEEKK 1031
Qy	334 DTPREIMPAPKETSFKTWAAGRPKIAWEKTPVTKGCVARVTSNKTIVLEKGRSKM 393
Db	1032 -LELNIQDLEEQLEEEBAARQKLEKVLQDAKIKKYBEDLALITDDQNKLLKEK---KL 1087
Qy	394 IACPTKSSSTKASANDQFPSEKOEDEEYSCDSRSLFESSAKIQVCIPESIQKVMEI 453
Db	1088 L-----ERANDL---SOTLAESEEEK-----AKHLAKLAKKHEATITE-LEERLHKD 1130
Qy	454 NREVEPPKPSAFKPAIEMQNSVFNKAPFLK---NEQTLRADPMFPPESEKOKOYEENSW 510
Db	1131 QOQKQESDRS-----KKEKTEVADLKEQLNERRVQVDEMQALAKEE----- 1173

1075 KYFEDIKILKEKN-----AELQMTLKLKEESLTKRASQYSQGLKVLIAE-----NTM 1122

1640 QLTEDAKLELVNMQALRSQFERDLLAKEEAGAEKRRGLVQLRDLTETELDEERKQRTAA 1699

1122 LTSKLKEKODKETEIAEIESHH-----PRLASAVQDHDQIQTVTRSK-----S 1162

1700 VASKKLEGLDELKEIETTMENHKVEDALKHAKLQAOVKDALRDAEAAKAAKEELQALS 1759

1163 QEPAPFHIAGDACLQRKXNVDSVT-----IYNNEVLHQPLSEARQKS-----KSLK 1208

1760 KEADGKVKALEAEVLQLTEDLASSERARRAAETERDELAEEIANNANKGSLMIDEKRRL 1819

1209 INLNYAGDALRENTLVSEHAQRDQRETQCOMKEAHHMYQNEQDNVNHK-----TEQOE 1261

1820 ARTATLEELSEEEQSNSEVLLDRSRQAQIQIBOLITTELANEKSQKNENGRRALLBRQNK 1879

1262 SLQDKLFLQQS-----KNMWLQOOLVHAHKKADNKS 1293

1880 ELKAKLAEIETAQRTKVKATIATLEAKIAKVEOLENEGKERLLOQK---ANRKMDDKIK 1936

1294 -ITIDHFLERKMQHLLKEKNEEIEFNYNHLLKNRIYQVEKE--KAETE 1339

1937 ELTMNIEDERRHVDQH--KEQMDKLNRSIKLLKRNLDTEEELOKEKTKQ 1983

Search completed: February 19, 2005, 02:12:17

Job time : 49.9263 secs

Search completed: February 19, 2005, 02:12:17
Job time : 49.9263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 02:11:01 ; Search time 125.953 Seconds
(without alignments)
3484.083 Million cell updates/sec

Title: US-09-602-362E-23

Perfect score: 6920

Sequence: 1 MTRKKKTINLNIDQAQRKA.....NHLKNRIYQYEKKAETENS 1341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6920	100.0	1341	13	US-10-007-805-565
2	6920	100.0	1341	14	US-10-076-622-565
3	6920	100.0	1341	14	US-10-177-293-334
4	6920	100.0	1341	14	US-10-124-805-565
5	6920	100.0	1341	15	US-10-058-270A-4
6	6892	99.6	1349	13	US-10-007-805-573
7	6892	99.6	1349	14	US-10-076-622-573
8	6892	99.6	1349	14	US-10-124-805-573
9	5130	74.1	1002	9	US-09-604-287A-475
10	5130	74.1	1002	9	US-09-834-759-475
11	5130	74.1	1002	10	US-09-551-621-475
12	5130	74.1	1002	13	US-10-007-805-475
13	5130	74.1	1002	14	US-10-076-622-475

14	5130	74.1	1002	14	US-10-124-805-475	Sequence 475, App
15	5130	74.1	1095	9	US-09-834-759-493	Sequence 493, App
16	5130	74.1	1095	13	US-10-007-805-493	Sequence 493, App
17	5130	74.1	1095	14	US-10-076-622-493	Sequence 493, App
18	5130	74.1	1095	14	US-10-124-805-493	Sequence 493, App
19	5124	74.0	1013	13	US-10-007-805-553	Sequence 553, App
20	5124	74.0	1013	14	US-10-076-622-553	Sequence 553, App
21	5124	74.0	1013	14	US-10-124-805-553	Sequence 553, App
22	3945	57.0	1239	13	US-10-007-805-577	Sequence 577, App
23	3945	57.0	1239	14	US-10-076-622-577	Sequence 577, App
24	3945	57.0	1239	14	US-10-124-805-577	Sequence 577, App
25	3721	53.8	1225	14	US-10-177-293-332	Sequence 332, App
26	3256.5	47.1	661	13	US-10-007-805-552	Sequence 552, App
27	3256.5	47.1	661	14	US-10-076-622-552	Sequence 552, App
28	3252.5	47.0	650	9	US-09-825-301-25	Sequence 25, Appl
29	3252.5	47.0	650	9	US-09-604-287A-469	Sequence 469, App
30	3252.5	47.0	650	9	US-09-834-759-469	Sequence 469, App
31	3252.5	47.0	650	10	US-09-551-621-469	Sequence 469, App
32	3252.5	47.0	650	13	US-10-007-805-494	Sequence 494, App
33	3252.5	47.0	743	13	US-10-076-622-494	Sequence 494, App
34	3252.5	47.0	743	14	US-10-124-805-494	Sequence 494, App
35	3252.5	47.0	650	14	US-10-124-805-469	Sequence 469, App
36	3252.5	47.0	650	14	US-10-033-527-25	Sequence 25, Appl
37	3252.5	47.0	743	9	US-09-834-759-494	Sequence 494, App
38	3252.5	47.0	743	13	US-10-007-805-494	Sequence 494, App
39	3252.5	47.0	743	14	US-10-124-805-494	Sequence 494, App
40	3252.5	47.0	743	14	US-10-124-805-494	Sequence 494, App
41	3075	44.4	1011	16	US-10-408-765A-1557	Sequence 1557, App
42	2589	37.4	512	16	US-10-181-663-16	Sequence 16, Appl
43	2270	32.8	466	9	US-09-825-301-28	Sequence 28, Appl
44	2270	32.8	466	9	US-09-604-287A-472	Sequence 472, App
45	2270	32.8	466	9	US-09-834-759-472	Sequence 472, App

ALIGNMENTS

RESULT 1

US-10-007-805-565
; Sequence 565, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margerita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 565
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-565

Query Match 100.0%; Score 6920; DB 13; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTRKKKTINLNIDQAQRKATLHWACVNGHEBVTFLVDRKCOLDVLDGEHRTPLMKALQC 60
Db 1 MTRKKKTINLNIDQAQRKATLHWACVNGHEBVTFLVDRKCOLDVLDGEHRTPLMKALQC 60

Qy	61	HOEACANILDSGADINLVDVYGNMALHYAVYSEILSVAKLLSHGAVIEVHNKASLTPL	120
Db	61	HOEACANILDSGADINLVDVYGNMALHYAVYSEILSVAKLLSHGAVIEVHNKASLTPL	120
Qy	121	LLSITKSEQIVFELLIKNANANAVKYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI	180
Db	121	LLSITKSEQIVFELLIKNANANAVKYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI	180
Qy	181	CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAALPAERTPTDAE	240
Db	181	CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAALPAERTPTDAE	240
Qy	241	SLVEKTPDEAALPAERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQSAEE	300
Db	241	SLVEKTPDEAALPAERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQSAEE	300
Qy	301	TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPRK	360
Db	301	TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPRK	360
Qy	361	IWEKKETPVKTGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSESQEE	420
Db	361	IWEKKETPVKTGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSESQEE	420
Qy	421	DEEYSCDSRSLFESSAKIQVCIPESIQYKVMINEVEEPPKPSAFKPAIEMQNSVFNK	480
Db	421	DEEYSCDSRSLFESSAKIQVCIPESIQYKVMINEVEEPPKPSAFKPAIEMQNSVFNK	480
Qy	481	AFELKNEOTLRADPMFPPESSKQDYEENSWSDESICETVSKQDVCLPKATHQKEIDKING	540
Db	481	AFELKNEOTLRADPMFPPESSKQDYEENSWSDESICETVSKQDVCLPKATHQKEIDKING	540
Qy	541	KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPEKPSAFEPATEMOKSVFNKA	600
Db	541	KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPEKPSAFEPATEMOKSVFNKA	600
Qy	601	LELKNQOTLRADIELPSESQKDYESSWDSLSLCTVSKQDVCLPKAAHQKEIDKINGK	660
Db	601	LELKNQOTLRADIELPSESQKDYESSWDSLSLCTVSKQDVCLPKAAHQKEIDKINGK	660
Qy	661	LEGSPVKDGLLKANGCMKVISIPTKALELMDMTFKAEPPEKPSAFEPATEMOKSVFNKAL	720
Db	661	LEGSPVKDGLLKANGCMKVISIPTKALELMDMTFKAEPPEKPSAFEPATEMOKSVFNKAL	720
Qy	721	ELKNQOTLRADIELPSESQKDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKL	780
Db	721	ELKNQOTLRADIELPSESQKDYESSWDSLSLCTVSKQDVCLPKATHQKEIDKINGKL	780
Qy	781	ESSPNDGFLKAPCRMKVISIPTKALELMDMTFKAEPPEKPSAFEPATEMOKSVFNKALE	840
Db	781	ESSPNDGFLKAPCRMKVISIPTKALELMDMTFKAEPPEKPSAFEPATEMOKSVFNKALE	840
Qy	841	LKNEOTLRADQMPFPPSESQKDYESSWDSLSLCTVSKQDVCPKATHQKEMDKISGKLE	900
Db	841	LKNEOTLRADQMPFPPSESQKDYESSWDSLSLCTVSKQDVCPKATHQKEMDKISGKLE	900
Qy	901	DSTSLSKILDTVHSCERABELOKHCEQRTGKMEQMKKFCVLLKKLSEAKKIKSOLQENQ	960
Db	901	DSTSLSKILDTVHSCERABELOKHCEQRTGKMEQMKKFCVLLKKLSEAKKIKSOLQENQ	960
Qy	961	KVKWEQELCSVRLTLNOEKEKRNADIINEKIREELGRIEEQRKLEVKQOLEQALRIQ	1020
Db	961	KVKWEQELCSVRLTLNOEKEKRNADIINEKIREELGRIEEQRKLEVKQOLEQALRIQ	1020
Qy	1021	DIELKSVESNLNQVSHTHENENYLLHENCWMLKKEIAMLKLEIATLKHQOEBKENKYPEDI	1080
Db	1021	DIELKSVESNLNQVSHTHENENYLLHENCWMLKKEIAMLKLEIATLKHQOEBKENKYPEDI	1080
Qy	1081	KILKEKNAELQWTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKLEKQDKEILEABIE	1140
Db	1081	KILKEKNAELQWTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKLEKQDKEILEABIE	1140
Qy	1141	SHHPRLASAVQDHQDIQVTSRKQEPAFHIAQDACLQRKKNVDVSVSTIYNNVHLHQPLSEA	1200
Db	1141	SHHPRLASAVQDHQDIQVTSRKQEPAFHIAQDACLQRKKNVDVSVSTIYNNVHLHQPLSEA	1200
Qy	1201	QRKSKSLKINLNYAGDALRENTLVSEHAORDORETOCQMKBAEHMYQNEQDNVKNKHEEQ	1260
Db	1201	QRKSKSLKINLNYAGDALRENTLVSEHAORDORETOCQMKBAEHMYQNEQDNVKNKHEEQ	1260
Qy	1261	ESLDQKLFQLOSQNMVLQOOLVHAHKKADNKSITIDHFLERKMOHLLKEKNEEIFNY	1320
Db	1261	ESLDQKLFQLOSQNMVLQOOLVHAHKKADNKSITIDHFLERKMOHLLKEKNEEIFNY	1320
Qy	1321	NNHLKNRIYQYEKEKAETENS 1341	
Db	1321	NNHLKNRIYQYEKEKAETENS 1341	

RESULT 2

US-10-076-622-565
 ; Sequence 565, Application US/10076622
 ; Publication No. US20030023036A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: Sleath, Paul R.
 ; APPLICANT: Persing, David H.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.470C11
 ; CURRENT APPLICATION NUMBER: US/10/076,622
 ; CURRENT FILING DATE: 2002-02-13
 ; NUMBER OF SEQ ID NOS: 627
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 565
 ; LENGTH: 1341
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-076-622-565

Query Match 100.0%; Score 6920; DB 14; Length 1341;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MTREKKTINLNIQDAQKRTALHWA	60
Db	1	MTREKKTINLNIQDAQKRTALHWA	60
Qy	61	HOEACANILDSGADINLVDVYGNMALHYAVYSEILSVAKLLSHGAVIEVHNKASLTPL	120
Db	61	HOEACANILDSGADINLVDVYGNMALHYAVYSEILSVAKLLSHGAVIEVHNKASLTPL	120
Qy	121	LLSITKSEQIVFELLIKNANANAVKYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI	180
Db	121	LLSITKSEQIVFELLIKNANANAVKYKCTALMLAVCHGSSEIVGMLLOQNVDFPAADI	180
Qy	181	CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAALPAERTPTDAE	240
Db	181	CGVTAHYAVTCGPHHIEHQIMEYIRKLSKNHONTNPEGTSAGTPDEAALPAERTPTDAE	240
Qy	241	SLVEKTPDEAALPAERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQSAEE	300
Db	241	SLVEKTPDEAALPAERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQSAEE	300
Qy	301	TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPRK	360
Db	301	TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPRK	360
Qy	361	IWEKKETPVKTGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSESQEE	420
Db	361	IWEKKETPVKTGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORFPSESQEE	420
Qy	421	DEEYSCDSRSLFESSAKIQVCIPESIQYKVMINEVEEPPKPSAFKPAIEMQNSVFNK	480
Db	421	DEEYSCDSRSLFESSAKIQVCIPESIQYKVMINEVEEPPKPSAFKPAIEMQNSVFNK	480

QY 481 APELKNQOTLRADPMPFPSPESKQDYENSWSSESICETVSKQDVCLPKATHQKEIDKING 540
DB 481 APELKNQOTLRADPMPFPSPESKQDYENSWSSESICETVSKQDVCLPKATHQKEIDKING 540
QY 541 KLEESPNKDGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKA 600
DB 541 KLEESPNKDGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKA 600
QY 601 LELKNEQOTRADEILPSESQKDYENSWSSESICETVSKQDVCLPKAAHQKEIDKINGK 660
DB 601 LELKNEQOTRADEILPSESQKDYENSWSSESICETVSKQDVCLPKAAHQKEIDKINGK 660
QY 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKAL 720
DB 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKAL 720
QY 721 ELKNEQOTLRADPMPFPSPESKQDYENSWSSESICETVSKQDVCLPKATHQKEIDKINGKL 780
DB 721 ELKNEQOTLRADPMPFPSPESKQDYENSWSSESICETVSKQDVCLPKATHQKEIDKINGKL 780
QY 781 BESPDNDGFLKAPCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALE 840
DB 781 BESPDNDGFLKAPCMKVSIPTKALELMDMOTFKABPPKPSAFEPATEMOKSVPNKALE 840
QY 841 LKNEQOTLRADPMPFPSPESKQDYENSWSSESICETVSKQDVCLPKATHQKEIDKINGKLE 900
DB 841 LKNEQOTLRADPMPFPSPESKQDYENSWSSESICETVSKQDVCLPKATHQKEIDKINGKLE 900
QY 901 DSTLSKILDTVHSCERARELOKHCEQRTGMEQMKKFCVLKKLSEAKEIKSOLBNQ 960
DB 901 DSTLSKILDTVHSCERARELOKHCEQRTGMEQMKKFCVLKKLSEAKEIKSOLBNQ 960
QY 961 KVKWEQELCSVRLTLNQEERKRNADILNEKIREELGRIEEORHKELEVKQOLEQALRIQ 1020
DB 961 KVKWEQELCSVRLTLNQEERKRNADILNEKIREELGRIEEORHKELEVKQOLEQALRIQ 1020
QY 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIATMLKLEIATLKHQYQKENKYFEDI 1080
DB 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIATMLKLEIATLKHQYQKENKYFEDI 1080
QY 1081 KILKEKNAELQWTLKKEESITKTRASQSGQLKVLIATENTMLTSKLKEQDKLEIABIE 1140
DB 1081 KILKEKNAELQWTLKKEESITKTRASQSGQLKVLIATENTMLTSKLKEQDKLEIABIE 1140
QY 1141 SHHPRLASAVODHDOI VTSRSQSEPAFHAGDACLQRMNVDSSTIYNNEVLHQPLSEA 1200
DB 1141 SHHPRLASAVODHDOI VTSRSQSEPAFHAGDACLQRMNVDSSTIYNNEVLHQPLSEA 1200
QY 1201 QRKSKSLINLYAGDALRENTLVSEHAQORDQRETQCMKEAEMHYQNEQDNVKNHTEBQ 1260
DB 1201 QRKSKSLINLYAGDALRENTLVSEHAQORDQRETQCMKEAEMHYQNEQDNVKNHTEBQ 1260
QY 1261 ESLDOKLQLOSKNNWLOQLVHAHKADNKSKITIDHFLERKQHLLKKEKNEEIPNY 1320
DB 1261 ESLDOKLQLOSKNNWLOQLVHAHKADNKSKITIDHFLERKQHLLKKEKNEEIPNY 1320
QY 1321 NNHLKNRIYQYEKEKAETENS 1341
DB 1321 NNHLKNRIYQYEKEKAETENS 1341

RESULT 3
US-10-177-293-334
; Sequence 334, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen

; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sabin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 334
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-293-334

Query Match 100.0%; Score 6920; DB 14; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTKRKKITNLNIQDAQKRTALHWACVNGHEEVTFLVDRKCOLDVLDGEHRTPLMKALQC 60
DB 1 MTKRKKITNLNIQDAQKRTALHWACVNGHEEVTFLVDRKCOLDVLDGEHRTPLMKALQC 60
QY 61 HOECANILIDSGADINLVVYGNMALHYAVYSEILSVVAKLHSHGAVIEVHNKASLTPL 120
DB 61 HOECANILIDSGADINLVVYGNMALHYAVYSEILSVVAKLHSHGAVIEVHNKASLTPL 120
QY 121 LLSITKRSQEVIFELLKNNANANAVNKYKCTALMLAVCHGSGSEIVGMLLQNVDFVFAADI 180
DB 121 LLSITKRSQEVIFELLKNNANANAVNKYKCTALMLAVCHGSGSEIVGMLLQNVDFVFAADI 180
QY 181 CGVTAHYAVTCGPHHIEHQIMEYIRKLSKHQNTNPGTSAGTPDEAPLAERTPDTA 240
DB 181 CGVTAHYAVTCGPHHIEHQIMEYIRKLSKHQNTNPGTSAGTPDEAPLAERTPDTA 240
QY 241 SLVEKTPDEAPLVERTPDTAESLVEKTPDEAASLVEGTSKIOCLEKATSGKPFQSAE 300
DB 241 SLVEKTPDEAPLVERTPDTAESLVEKTPDEAASLVEGTSKIOCLEKATSGKPFQSAE 300
QY 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEEDTPREIMSPAKETSEKFTWAAKGRPRK 360
DB 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEEDTPREIMSPAKETSEKFTWAAKGRPRK 360
QY 361 IAWKKEETPVKTCGVARTSVNKTVLEKGRSKMIACPTKESSTKASANDORPPSPESKQEE 420
DB 361 IAWKKEETPVKTCGVARTSVNKTVLEKGRSKMIACPTKESSTKASANDORPPSPESKQEE 420
QY 421 DEEYSCDSRSLFESSAKIQCIPESYOKVMEINREVEEPPKPSAFKPAIEMQNSVFNK 480
DB 421 DEEYSCDSRSLFESSAKIQCIPESYOKVMEINREVEEPPKPSAFKPAIEMQNSVFNK 480

QY	481	AFELKNEQTLRADPMFPPESSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKING	540
DB	481	AFELKNEQTLRADPMFPPESSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKING	540
QY	541	KLESPNKDGLLKATCGMKVSIPTKALELMDQTFKAEPPGKPSAFEPATEMOKSVNKA	600
DB	541	KLESPNKDGLLKATCGMKVSIPTKALELMDQTFKAEPPGKPSAFEPATEMOKSVNKA	600
QY	601	LELNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKAAHQKEIDKINGK	660
DB	601	LELNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKAAHQKEIDKINGK	660
QY	661	LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	720
DB	661	LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	720
QY	721	ELKNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKATHQKEIDKINGKL	780
DB	721	ELKNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKATHQKEIDKINGKL	780
QY	781	BESPDNDGFLKAPCRMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	840
DB	781	BESPDNDGFLKAPCRMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	840
QY	841	LKNEQTLRADQMFPSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKINGK	900
DB	841	LKNEQTLRADQMFPSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKINGK	900
QY	901	DSTLSKILDTVHSCERARELQDCEQRTGRMEQMKKFCVYLKCKLSEAKBIKSQLENQ	960
DB	901	DSTLSKILDTVHSCERARELQDCEQRTGRMEQMKKFCVYLKCKLSEAKBIKSQLENQ	960
QY	961	KVWQOELCSVRLTLNOBEKERNADILNEKIREELGRIEBOHRKELEVKQOELALRIQ	1020
DB	961	KVWQOELCSVRLTLNOBEKERNADILNEKIREELGRIEBOHRKELEVKQOELALRIQ	1020
QY	1021	DIELKSVESNLQVSTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQENKIFYEDI	1080
DB	1021	DIELKSVESNLQVSTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQENKIFYEDI	1080
QY	1081	KILKERNABLOMTLKLKESSLTKRASQYSGQLKVLAIENTMTLSKLKQKQKEILEABIE	1140
DB	1081	KILKERNABLOMTLKLKESSLTKRASQYSGQLKVLAIENTMTLSKLKQKQKEILEABIE	1140
QY	1141	SHHPRLASVQDHQIVTSRKQOEPAFHAGDACLQKKNVDVSTIYNNEVLHOPLSEA	1200
DB	1141	SHHPRLASVQDHQIVTSRKQOEPAFHAGDACLQKKNVDVSTIYNNEVLHOPLSEA	1200
QY	1201	QRKSKSLKINLVAGDALRENTLVSEHAQDQRETQCKMEAHMYQNEQDNNVKNHTEQQ	1260
DB	1201	QRKSKSLKINLVAGDALRENTLVSEHAQDQRETQCKMEAHMYQNEQDNNVKNHTEQQ	1260
QY	1261	ESLDQKLFLQSKNMMLQOOLVHAHKADNKSITIDIHFLERKMOHLLKEKNEIFNY	1320
DB	1261	ESLDQKLFLQSKNMMLQOOLVHAHKADNKSITIDIHFLERKMOHLLKEKNEIFNY	1320
QY	1321	NNHLKNRIQYKEKEAETENS	1341
DB	1321	NNHLKNRIQYKEKEAETENS	1341

RESULT 4

US-10-124-805-565

; Sequence 565, Application US/10124805

; Publication No. US20030166022A1

; GENERAL INFORMATION:

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Persing, David H.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C12

; CURRENT APPLICATION NUMBER: US/10/124,805

QY	1	MTKRKKTINLINIQAOKRTALHWACVNGHEEVVTVFLVDRKCCQLDVLDSHRTPLMKALOC	60
DB	1	MTKRKKTINLINIQAOKRTALHWACVNGHEEVVTVFLVDRKCCQLDVLDSHRTPLMKALOC	60
QY	61	HOACANILIDSGADINLVDTVGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL	120
DB	61	HOACANILIDSGADINLVDTVGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL	120
QY	121	LLSITTKRSEQIVFELIKNANANANVYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI	180
DB	121	LLSITTKRSEQIVFELIKNANANANVYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI	180
QY	181	CGVTABHYAVTCGFHHIHEQIMEYIRKLSKNHONTNPEGTSAGTDPDEAALPABRTPTAE	240
DB	181	CGVTABHYAVTCGFHHIHEQIMEYIRKLSKNHONTNPEGTSAGTDPDEAALPABRTPTAE	240
QY	241	SLVEKTPDEAALPLVERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGFPEOSAE	300
DB	241	SLVEKTPDEAALPLVERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGFPEOSAE	300
QY	301	TPREITSPAKETSEKFTWPAKGRPRIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPRK	360
DB	301	TPREITSPAKETSEKFTWPAKGRPRIAWEKKEDTTPREIMSPAKETSEKFTWAAKGRPRK	360
QY	361	IAMEKTPVKTGCVARVTSNKTVLEKGRSKMIACPKTESSTKASANDORFFSESKQEE	420
DB	361	IAMEKTPVKTGCVARVTSNKTVLEKGRSKMIACPKTESSTKASANDORFFSESKQEE	420
QY	421	DEEYSCDSRSIFESSAKIQVCIPESIIQKWEINREVEEPPKPSAFKPAIEMQNSVPNK	480
DB	421	DEEYSCDSRSIFESSAKIQVCIPESIIQKWEINREVEEPPKPSAFKPAIEMQNSVPNK	480
QY	481	AFELKNEQTLRADPMFPPESSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKING	540
DB	481	AFELKNEQTLRADPMFPPESSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKING	540
QY	541	KLESPNKDGLLKATCGMKVSIPTKALELMDQTFKAEPPGKPSAFEPATEMOKSVNKA	600
DB	541	KLESPNKDGLLKATCGMKVSIPTKALELMDQTFKAEPPGKPSAFEPATEMOKSVNKA	600
QY	601	LELNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKAAHQKEIDKINGK	660
DB	601	LELNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKAAHQKEIDKINGK	660
QY	661	LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	720
DB	661	LEGSPVKDGLLKANGCMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	720
QY	721	ELKNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKATHQKEIDKINGKL	780
DB	721	ELKNEQTLRADEILPSESKQDYENSWDTSSELCTVSKQDVCLPKATHQKEIDKINGKL	780
QY	781	BESPDNDGFLKAPCRMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	840
DB	781	BESPDNDGFLKAPCRMKVISIPTKALELMDQTFKAEPPPEKPSAFEPATEMOKSVPNKA	840
QY	841	LKNEQTLRADQMFPSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKINGK	900
DB	841	LKNEQTLRADQMFPSKQDYENSWDSSELCTVSKQDVCLPKATHQKEIDKINGK	900

Qy 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEOMKKKFCVLKKLSEAKEIKSOLENQ 960
Db 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEOMKKKFCVLKKLSEAKEIKSOLENQ 960
Qy 961 KVKWEQELCSVRLTLNQEEERKNADILNEKIREELGRIEBOHRKELEVKQOLQALRIQ 1020
Db 961 KVKWEQELCSVRLTLNQEEERKNADILNEKIREELGRIEBOHRKELEVKQOLQALRIQ 1020
Qy 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSKENKYPEDI 1080
Db 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSKENKYPEDI 1080
Qy 1081 KILKEKNAELQMTLKLKEESITKASQYSGQLKVLIAENTMLTSKLKEKQDEILEABIE 1140
Db 1081 KILKEKNAELQMTLKLKEESITKASQYSGQLKVLIAENTMLTSKLKEKQDEILEABIE 1140
Qy 1141 SHHPRLASAVODHQIVTSRSQBPAPHIAGDACLQRMNVDSSTIYNNVHLQPLSEA 1200
Db 1141 SHHPRLASAVODHQIVTSRSQBPAPHIAGDACLQRMNVDSSTIYNNVHLQPLSEA 1200
Qy 1201 QRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAHEMYQNEQDNVKNHTEQQ 1260
Db 1201 QRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAHEMYQNEQDNVKNHTEQQ 1260
Qy 1261 ESLQKLPOLQSKNMWLOQQLVHAHKKADNKSKITIDHFLERKQMHLLKEKNEEIPNY 1320
Db 1261 ESLQKLPOLQSKNMWLOQQLVHAHKKADNKSKITIDHFLERKQMHLLKEKNEEIPNY 1320
Qy 1321 NNHLKNRIYQYEKAEKATENS 1341
Db 1321 NNHLKNRIYQYEKAEKATENS 1341

RESULT 5

US-10-058-270A-4
; Sequence 4, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Afar, Daniel
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and
; FILE REFERENCE: 018501-005210US
; CURRENT APPLICATION NUMBER: US/10/058,270A
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,965
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/265,928
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 09/829,472
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/282,698
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/288,590
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,443
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 4
; LENGTH: 1341
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-270A-4

Query Match 100.0%; Score 6920; DB 15; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKKKTINLNQDAQKRTALHWCVNGHEBVTFLVDRKCQLDVLGDEHRTPLMKALQC 60
|||||

Db 1 MTKKKTINLNQDAQKRTALHWCVNGHEBVTFLVDRKCQLDVLGDEHRTPLMKALQC 60
Qy 61 HQEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLSHGAVIEVHNKASLTPL 120
Db 61 HQEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLSHGAVIEVHNKASLTPL 120
Qy 121 LLSITKSEQIVFELLIKNANANAVNYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180
Db 121 LLSITKSEQIVFELLIKNANANAVNYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180
Qy 181 CGVTAEHYAVTCGPHHHEQIMEYIRKLSKNHNTNPGTSAGTPDEAAPLAERTPDTAE 240
Db 181 CGVTAEHYAVTCGPHHHEQIMEYIRKLSKNHNTNPGTSAGTPDEAAPLAERTPDTAE 240
Qy 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVEGTSVKIQCLEKATSGKFEQSABE 300
Db 241 SLVEKTPDEAAPLVERTPDTAESLVEKTPDEAASLVEGTSVKIQCLEKATSGKFEQSABE 300
Qy 301 TPREITSPAKETSEKFTWPAKGRPRKTAWEKKEOTPREIMSPAKETSEKFTWAAKGRPRK 360
Db 301 TPREITSPAKETSEKFTWPAKGRPRKTAWEKKEOTPREIMSPAKETSEKFTWAAKGRPRK 360
Qy 361 IAWKKEKTPVKTGCVARVTSNKTVLEKGRSKMIACPTKESSTKASANDORFPSESKEE 420
Db 361 IAWKKEKTPVKTGCVARVTSNKTVLEKGRSKMIACPTKESSTKASANDORFPSESKEE 420
Qy 421 DEBYSCDSRSLFESSAKIQVCIPESIQKWEINREVEEPPKPSAFKPAIEMQNSVFNK 480
Db 421 DEBYSCDSRSLFESSAKIQVCIPESIQKWEINREVEEPPKPSAFKPAIEMQNSVFNK 480
Qy 481 APELKNQETLRADPMFPPESKQKDYEENSWSLCECTVSKQDVCLPKATHQKEIDKING 540
Db 481 APELKNQETLRADPMFPPESKQKDYEENSWSLCECTVSKQDVCLPKATHQKEIDKING 540
Qy 541 KLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKAEPPCKPSAFEPATEMOKSVPNKA 600
Db 541 KLEESPNKDGILLKATCGMKVSIPTKALELMDQTFKAEPPCKPSAFEPATEMOKSVPNKA 600
Qy 601 LELKNEQTRADEILPSESQKDYEENSWSLCECTVSKQDVCLPKAAHQKEIDKINGK 660
Db 601 LELKNEQTRADEILPSESQKDYEENSWSLCECTVSKQDVCLPKAAHQKEIDKINGK 660
Qy 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDQTFKAEPPCKPSAFEPATEMOKSVPNKAL 720
Db 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDQTFKAEPPCKPSAFEPATEMOKSVPNKAL 720
Qy 721 ELKNEQTLRADEILPSESQKDYEENSWSLCECTVSKQDVCLPKATHQKEIDKINGK 780
Db 721 ELKNEQTLRADEILPSESQKDYEENSWSLCECTVSKQDVCLPKATHQKEIDKINGK 780
Qy 781 EESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPCKPSAFEPATEMOKSVPNKALE 840
Db 781 EESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPCKPSAFEPATEMOKSVPNKALE 840
Qy 841 LKNEQTLRADQMFPSESQKDYEENSWSLCECTVSKQDVCLPKATHQKEIDKINGK 900
Db 841 LKNEQTLRADQMFPSESQKDYEENSWSLCECTVSKQDVCLPKATHQKEIDKINGK 900
Qy 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEOMKKKFCVLKKLSEAKEIKSOLENQ 960
Db 901 DSTSLSKILDTVHSCERARELOKHCEORTGMEOMKKKFCVLKKLSEAKEIKSOLENQ 960
Qy 961 KVKWEQELCSVRLTLNQEEERKNADILNEKIREELGRIEBOHRKELEVKQOLQALRIQ 1020
Db 961 KVKWEQELCSVRLTLNQEEERKNADILNEKIREELGRIEBOHRKELEVKQOLQALRIQ 1020
Qy 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSKENKYPEDI 1080
Db 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQSKENKYPEDI 1080
Qy 1081 KILKEKNAELQMTLKLKEESITKASQYSGQLKVLIAENTMLTSKLKEKQDEILEABIE 1140
Db 1081 KILKEKNAELQMTLKLKEESITKASQYSGQLKVLIAENTMLTSKLKEKQDEILEABIE 1140

Qy 1141 SHPRLASAVODHDOIVTSRKSQEPAFHAGDACLQRKMNVDVSSTIYNNEVLHOPLSEA 1200
Db 1141 SHPRLASAVODHDOIVTSRKSQEPAFHAGDACLQRKMNVDVSSTIYNNEVLHOPLSEA 1200
Qy 1201 QRKSKSLINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTEQ 1260
Db 1201 QRKSKSLINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTEQ 1260
Qy 1261 ESLDQKLFQLOSKNMWLOOQLVHAHKADNKSKITIDHFLERKQOHLLKEKNEEIFNY 1320
Db 1261 ESLDQKLFQLOSKNMWLOOQLVHAHKADNKSKITIDHFLERKQOHLLKEKNEEIFNY 1320
Qy 1321 NNHLKNRIYOYEKEKAETENS 1341
Db 1321 NNHLKNRIYOYEKEKAETENS 1341

RESULT 6
US-10-007-805-573
; Sequence 573, Application US/10007805
; Publication No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-573

Query Match 99.6%; Score 6892; DB 13; Length 1349;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TTKKKTININIQDAKRTALHWACVNGHEEVTVFLVDRKCOLVDLDGHRTPLMKALQCH 61
Db 10 TTKKKTININIQDAKRTALHWACVNGHEEVTVFLVDRKCOLVDLDGHRTPLMKALQCH 69
Qy 62 QACANILDSGADINLVGVGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPLL 121
Db 70 QACANILDSGADINLVGVGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPLL 129
Qy 122 LSITRSEQIVFLLIKNANANVYKCTALMLAVCHGSSBIVGMILQONVDVFAADIC 181
Db 130 LSITRSEQIVFLLIKNANANVYKCTALMLAVCHGSLBIVGMILQONVDVFAADIC 189
Qy 182 GYTAEHYAVTCGFHHIHEQIMYIRKLSKNHONTPEGTSAGTDPDEAAPLAERTPTDAES 241
Db 190 GYTAEHYAVTCGFHHIHEQIMYIRKLSKNHONTPEGTSAGTDPDEAAPLAERTPTDAES 249
Qy 242 LVEKTPDEAAPLVERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQASBET 301
Db 250 LVEKTPDEAAPLVERTPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQASBET 309
Qy 302 PREITSPAKETSEKFTWPAKGRPRKIATWEKKEDTPREIMSPAKETSEKFTWPAKGRPRKI 361

RESULT 7

US-10-076-622-573

```
; Sequence 573, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-076-622-573

Query Match      99.6%; Score 6892; DB 14; Length 1349;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKQCOLDVLDGEHRTPLMKALQCH 61
Db 10 TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKQCOLDVLDGEHRTPLMKALQCH 69
Qy 62 QEACANILIDSGADINLVGVGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 121
Db 70 QEACANILIDSGADINLVGVGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 129
Qy 122 LSITKRSEQIVFELLKNNANANVKNYKCTALMLAVCHSGSSEIVGMLLQNVVFAADIC 181
Db 130 LSITKRSEQIVFELLKNNANANVKNYKCTALMLAVCHSGSSEIVGMLLQNVVFAADIC 189
Qy 182 GVTAEHAVTCGPHIHHQMEYIRKLSKNHNTNPEGTSAGTDEAPLAERTPTDAES 241
Db 190 GVTAEHAVTCGPHIHHQMEYIRKLSKNHNTNPEGTSAGTDEAPLAERTPTDAES 249
Qy 242 LVEKTPDEAAPLVERTPTDAESLVKTPDEASLVGTSKIQCLEKATSGKFPQSABET 301
Db 250 LVEKTPDEAAPLVERTPTDAESLVKTPDEASLVGTSKIQCLEKATSGKFPQSABET 309
Qy 302 PREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRKI 361
Db 310 PREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRKI 369
Qy 362 AWEKKEPTVKTGCVARVTSNKTIVLEKGRSWMIACTPKESSTKASANDQRPSPSSQBED 421
Db 370 AWEKKEPTVKTGCVARVTSNKTIVLEKGRSWMIACTPKESSTKASANDQRPSPSSQBED 429
Qy 422 BEYSCDSRSLPSSAKIQVCIPESIQYKVMIEINREVEEPPKPSAFKPAIMQNSVPNKA 481
Db 430 BEYSCDSRSLPSSAKIQVCIPESIQYKVMIEINREVEEPPKPSAFKPAIMQNSVPNKA 489
Qy 482 FELKNEQTLRADPMPFPSSKQKDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGK 541
Db 490 FELKNEQTLRADPMPFPSSKQKDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGK 549
Qy 542 LEESPNDGGLLKATCGMKVSIPTKALELMDMTQTFKAEPPKPSAFEPATENQKSVPNKAL 601
Db 550 LEESPNDGGLLKATCGMKVSIPTKALELMDMTQTFKAEPPKPSAFEPATENQKSVPNKAL 609
Qy 602 ELKNEQTLRADEILPSSKQKDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGK 661
Db 610 ELKNEQTLRADEILPSSKQKDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGK 669
Qy 662 EGSPVKDGLLKANGCMKVISIPTKALELMDMTQTFKAEPPKPSAFEPATENQKSVPNKALE 721
Db 670 EGSPVKDGLLKANGCMKVISIPTKALELMDMTQTFKAEPPKPSAFEPATENQKSVPNKALE 729
Qy 722 LKNEQTLRADBILPSSKQKDYESSWSDESICETVSQKDVCLPKATHQKEIDKINGKLE 781
Db 730 LKNEQTLRADBILPSSKQKDYESSWSDESICETVSQKDVCLPKATHQKEIDKINGKLE 789
Qy 782 ESPNDGGLLKAPCRMKVISIPTKALELMDMTQTFKAEPPKPSAFEPATENQKSVPNKALEL 841
Db 790 ESPNDGGLLKAPCRMKVISIPTKALELMDMTQTFKAEPPKPSAFEPATENQKSVPNKALEL 849
Qy 842 KNEQTLRADQMPFPSSKQKDYENSWSDESILRETVSQKDVCPKATHQKEMDKISGKLE 901
Db 850 KNEQTLRADQMPFPSSKQKDYENSWSDESILRETVSQKDVCPKATHQKEMDKISGKLE 909
Qy 902 STLSKILDTVHSCERARELQKHCEORTGMEQMKKKFCVLKKLSEAKIKSQLENOK 961
Db 910 STLSKILDTVHSCERARELQKHCEORTGMEQMKKKFCVLKKLSEAKIKSQLENOK 969
Qy 962 VKWEQELCSVRLTLNQBEERKRNADILNEKIRELGRIEEQHREKELEVKQOLEQALRIQD 1021
Db 970 VKWEQELCSVRLTLNQBEERKRNADILNEKIRELGRIEEQHREKELEVKQOLEQALRIQD 1029
Qy 1022 IELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIK 1081
Db 1030 IELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIK 1089
Qy 1082 ILKEKNAELQWTLKKEESLTKRASQYSGQLKVLIAENTMLTSLKKEKQDKKEILEAEIES 1141
Db 1090 ILKEKNAELQWTLKKEESLTKRASQYSGQLKVLIAENTMLTSLKKEKQDKKEILEAEIES 1149
Qy 1142 HHPLASAVODHDIQVTSRKSQEPAFHIIAGDACLQRKMNVDVSTIYNNEVLHQPPLSAQ 1201
Db 1150 HHPLASAVODHDIQVTSRKSQEPAFHIIAGDACLQRKMNVDVSTIYNNEVLHQPPLSAQ 1209
Qy 1202 RKSLSKLINLYAGDALRENTLVSEHAQRDQRETQCMKEAHMYQNEQDNVNNKHTEQOE 1261
Db 1210 RKSLSKLINLYAGDALRENTLVSEHAQRDQRETQCMKEAHMYQNEQDNVNNKHTEQOE 1269
Qy 1262 SLDOQLFQLOSKNMLQOQLVHAHKADNKSITIDIFHLERKMOHLLKKEKBEIFNYN 1321
Db 1270 SLDOQLFQLOSKNMLQOQLVHAHKADNKSITIDIFHLERKMOHLLKKEKBEIFNYN 1329
Qy 1322 NHLKNRIYQYEKKAETE 1339
Db 1330 NHLKNRIYQYEKKAETE 1347

RESULT 8
US-10-124-805-573
; Sequence 573, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 573
; LENGTH: 1349
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-124-805-573

Query Match      99.6%; Score 6892; DB 14; Length 1349;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKQCOLDVLDGEHRTPLMKALQCH 61
Db 10 TKRKTINLNIDQAQKRTALHWACVNGHEEVVTFVDRKQCOLDVLDGEHRTPLMKALQCH 69
Qy 62 QEACANILIDSGADINLVGVGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 121
```

Db 70 QEACANILLISGADINLVDTVGNMALHAYVSEILSVAKLLSHGAVIEVHKNASLTPLL 129
Qy 122 LSITKRSQIIVFLLIKNANANAVNKYKCTALMLAVCHGSSEIVGMLLQQNVVFAADIC 181
Db 130 LSITKRSQIIVFLLIKNANANAVNKYKCTALMLAVCHGLSEIVGMLLQQNVVFAADIC 189
Qy 182 GVTAEHYAVTCGFPHIHQIIMEYIRKLSKNQNTNPEGSTAGTDEAAPLAERTPDTAES 241
Db 190 GVTAEHYAVTCGFPHIHQIIMEYIRKLSKNQNTNPEGSTAGTDEAAPLAERTPDTAES 249
Qy 242 LVEKTPDEAAPLVERTPDTASLSVEKTPDEAASIVEGTSDKIQCLEKATSKGFKQSAAET 301
Db 250 LVEKTPDEAAPLVERTPDTASLSVEKTPDEAASIVEGTSDKIQCLEKATSKGFKQSAAET 309
Qy 302 PREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRKI 361
Db 310 PREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRKI 369
Qy 362 AWEKKEPTVKTGCVARVTSNKTUVEKGRSKMIACPTKESSTKASANDORPPSKQBED 421
Db 370 AWEKKEPTVKTGCVARVTSNKTUVEKGRSKMIACPTKESSTKASANDORPPSKQBED 429
Qy 422 BEYSCDSRSLPSSAKIQVCIPESYQKWEINREVEPPKPSAFKPAIEMQNSVPNKA 481
Db 430 BEYSCDSRSLPSSAKIQVCIPESYQKWEINREVEPPKPSAFKPAIEMQNSVPNKA 489
Qy 482 FELKNEQTLRADMPMPPESSKQKDYENSWSDESLSCTVSQKDVCLPKATHQKEIDKINGK 541
Db 490 FELKNEQTLRADMPMPPESSKQKDYENSWSDESLSCTVSQKDVCLPKATHQKEIDKINGK 549
Qy 542 LEESPNDGGLLKATCGMKVSIPTKALELDMQTFKAEPPEKPSAFEPATEMOKSVPNKAL 601
Db 550 LEESPNDGGLLKATCGMKVSIPTKALELDMQTFKAEPPEKPSAFEPATEMOKSVPNKAL 609
Qy 602 ELKNEQTLRADMPPESSKQKDYENSWSDESLSCTVSQKDVCLPKATHQKEIDKINGK 661
Db 610 ELKNEQTLRADMPPESSKQKDYENSWSDESLSCTVSQKDVCLPKATHQKEIDKINGK 669
Qy 662 EGSPVKDGLLKANCQMKVSIPTKALELDMQTFKAEPPEKPSAFEPATEMOKSVPNKAL 721
Db 670 EGSPVKDGLLKANCQMKVSIPTKALELDMQTFKAEPPEKPSAFEPATEMOKSVPNKAL 729
Qy 722 LKNEQTLRADELTPSSKQKDYENSWSDESLSCTVSQKDVCLPKATHQKEIDKINGK 781
Db 730 LKNEQTLRADELTPSSKQKDYENSWSDESLSCTVSQKDVCLPKATHQKEIDKINGK 789
Qy 782 ESPDNDGFLKAPCRMVSIPTKALELDMQTFKAEPPEKPSAFEPATEMOKSVPNKAL 841
Db 790 ESPDNDGFLKAPCRMVSIPTKALELDMQTFKAEPPEKPSAFEPATEMOKSVPNKAL 849
Qy 842 KNEQTLRADMPPESSKQKDYENSWSDESLSCTVSQKDVCPKATHQKEMDKISGKLED 901
Db 850 KNEQTLRADMPPESSKQKDYENSWSDESLSCTVSQKDVCPKATHQKEMDKISGKLED 909
Qy 902 STSLSKILDTVHSCREARELOKHQBORTGKMEQMKKFCVLKXLSAKETKSOLENOK 961
Db 910 STSLSKILDTVHSCREARELOKHQBORTGKMEQMKKFCVLKXLSAKETKSOLENOK 969
Qy 962 VKWEQELCSVRLTNOQEEKRRNADILNEKIREELGRIBEOHQRKELEVQKQLEQALRIQD 1021
Db 970 VKWEQELCSVRLTNOQEEKRRNADILNEKIREELGRIBEOHQRKELEVQKQLEQALRIQD 1029
Qy 1022 IELKSVESNLNQVSHTHENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENKYPEDIK 1081
Db 1030 IELKSVESNLNQVSHTHENYLLHNCMLKKEIAMLKLEIATLKHQYOEKENKYPEDIK 1089
Qy 1082 ILKEKNAELQMTLKLKEESITKTRASQYSGOLKVLIAENTMLATSKLEKQDKLEILAEIES 1141
Db 1090 ILKEKNAELQMTLKLKEESITKTRASQYSGOLKVLIAENTMLATSKLEKQDKLEILAEIES 1149
Qy 1142 HHPRLASAVQDHQIIVTSRKSQEPAFHAGDACLQRKQNVDSSTIYNNNEVLHQPULSEAQ 1201

Db 1150 HHPRLASAVQDHQIIVTSRKSQEPAFHAGDACLQRKQNVDSSTIYNNNEVLHQPULSEAQ 1209
Qy 1202 RKSKSLKINLNYAGDALRENTLVSEHAQRDQRTQCKMEAEHMYQNEQDNVKNKHTEQOE 1261
Db 1210 RKSKSLKINLNYAGDALRENTLVSEHAQRDQRTQCKMEAEHMYQNEQDNVKNKHTEQOE 1269
Qy 1262 SLQDKLFLQSKQNMWLOQQLVHAHKADNKSKITIDHFLERKQVQHLLKKEKNEEIPFNYN 1321
Db 1270 SLQDKLFLQSKQNMWLOQQLVHAHKADNKSKITIDHFLERKQVQHLLKKEKNEEIPFNYN 1329
Qy 1322 NHLKNRIYQYEKKAETE 1339
Db 1330 NHLKNRIYQYEKKAETE 1347

RESULT 9

US-09-604-287A-475
; Sequence 475, Application US/09604287A
; Patent No. US20020064872A1

GENERAL INFORMATION:

; APPLICANT: Jiang, Yugui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER

; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604, 287A
; CURRENT FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 489

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 475

; LENGTH: 1002

; TYPE: PRT

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(1002)

; OTHER INFORMATION: Xaa = Any Amino Acid

; US-09-604-287A-475

Query Match 74.1%; Score 5130; DB 9; Length 1002;

Best Local Similarity 99.4%; Pred. No. 1.6e-257;

Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTGCVARVTSNKTUVEKGRSKMIACPTK 399

Db 1 MSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTGCVARVTSNKTUVEKGRSKMIACPTK 60

Qy 400 ESSTKASANDORPPSESKQBEDEYSQDSRSLPSSAKIQVCIPESYQKWEINREVEE 459

Db 61 ESSTKASANDORPPSESKQBEDEYSQDSRSLPSSAKIQVCIPESYQKWEINREVEE 120

Qy 460 PPKPSAFKPAIEMQNSVPNKAFLKNEQTLRADMPMPPESSKQKDYENSWSDESLSCTV 519

Db 121 PPKPSAFKPAIEMQNSVPNKAFLKNEQTLRADMPMPPESSKQKDYENSWSDESLSCTV 180

Qy 520 SQKDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELDMQTFKAEP 579

Db 181 SQKDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELDMQTFKAEP 240

Qy 580 PKGSAPFAPATEMOKSVPNKALELKNEQTLRADMPPESSKQKDYENSWSDESLSCTVS 639

Db 241 PKGSAPFAPATEMOKSVPNKALELKNEQTLRADMPPESSKQKDYENSWSDESLSCTVS 300

Qy 640 QKDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELDMQTFKAEP 699

Db 301 QKDVCLPKATHQKEIDKINGKLEESPNKQGLLKATCGMKVSIPTKALELDMQTFKAEP 360

Qy 700 EKPSAFEPATEMOKSVPNKALELKNEQTLRADMPPESSKQKDYENSWSDESLSCTVSQ 759

Db 361 EKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYESSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 819
Db 421 KDVCPLPKAXHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQDYESSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 939
Db 541 DVCVPKATHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 600
Qy 940 FCVLKKLSEAKEIKSQLENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKEIKSQLENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 BEQHRKELEVQKQLEQALRIQDIELKSVEESLNQVSHTHENENYLLHNCMLKKEIAMLK 1059
Db 661 BEQHRKELEVQKQLEQALRIQDIELKSVEESLNQVSHTHENENYLLHNCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESI-TKRASOYSGOLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESI-TKRASOYSGOLKVLIAEN 780
Qy 1120 TMLTSKLKEQDKLEIAEIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 1179
Db 781 TMLTSKLKEQDKLEIAEIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 840
Qy 1180 NVDVSTTYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 841 NVDVSTTYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 900
Qy 1240 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMMLQOQLVHAHKADNKSKITIDIH 1299
Db 901 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMMLQOQLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKQVHLLKKEKNEEIFNYYNNHLLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQVHLLKKEKNEEIFNYYNNHLLKNRIYQYEKEKAETENS 1002

RESULT 10

US-09-834-759-475
; Sequence 475, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William F.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-475

Query Match

74.1%; Score 5130; DB 9; Length 1002;

RESULT 11

Best Local Similarity 99.4%; Pred. No. 1.6e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 340 MSPAKETSEKETWAAGRPRIANEKKEETPVKTGCVARVTGNKTIVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKETWAAGRPRIANEKKEETPVKTGCVARVTGNKTIVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDQRPSESKQEEDEYSCDSRSLFESSAKIQVCIPESIQYKVMENREVEE 459
Db 61 ESSTKASANDQRPSESKQEEDEYSCDSRSLFESSAKIQVCIPESIQYKVMENREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAPFLKNEQTLRADQMPFSESKQDYESSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAPFLKNEQTLRADQMPFSESKQDYESSWDSLSLCTV 180
Qy 520 SQDVCPLPKATHQKIDKINGKLEESPNKOGLLKATCGMKVSIPTKALELMDQTFKAEP 579
Db 181 SQDVCPLPKATHQKIDKINGKLEESPNKOGLLKATCGMKVSIPTKALELMDQTFKAEP 240
Qy 580 PGKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYESSWDSLSLCTVS 639
Db 241 PGKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYESSWDSLSLCTVS 300
Qy 640 QKDVCPLPKAXHQKIDKINGKLEESPVKDGILLKANCGRMKVSIPTKALELMDQTFKAEP 699
Db 301 QKDVCPLPKAXHQKIDKINGKLEESPVKDGILLKANCGRMKVSIPTKALELMDQTFKAEP 360
Qy 700 EKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYESSWDSLSLCTVSQ 759
Db 361 EKPSAFPAIEMQKSVPNKALELKNQETLRADIELPSESKQDYESSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 819
Db 421 KDVCPLPKAXHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQDYESSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQETLRADQMPFSESKQDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 939
Db 541 DVCVPKATHQKIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPE 600
Qy 940 FCVLKKLSEAKEIKSQLENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKEIKSQLENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 BEQHRKELEVQKQLEQALRIQDIELKSVEESLNQVSHTHENENYLLHNCMLKKEIAMLK 1059
Db 661 BEQHRKELEVQKQLEQALRIQDIELKSVEESLNQVSHTHENENYLLHNCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESI-TKRASOYSGOLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKLKEESI-TKRASOYSGOLKVLIAEN 780
Qy 1120 TMLTSKLKEQDKLEIAEIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 1179
Db 781 TMLTSKLKEQDKLEIAEIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 840
Qy 1180 NVDVSTTYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 841 NVDVSTTYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 900
Qy 1240 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMMLQOQLVHAHKADNKSKITIDIH 1299
Db 901 KEAHEMYQNEODNVNKHTEQOESLDQKLFQOSKNMMLQOQLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKQVHLLKKEKNEEIFNYYNNHLLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQVHLLKKEKNEEIFNYYNNHLLKNRIYQYEKEKAETENS 1002


```
Db 121 PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPESSKQKQYENSWSSESICETV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 240
Qy 580 PKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 639
Db 241 PKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 360
Qy 700 EKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 759
Db 361 EKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 420
Qy 760 KDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 819
Db 421 KDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 480
Qy 820 KPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 879
Db 481 KPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 540
Qy 880 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 939
Db 541 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 600
Qy 940 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTNQEEKRRNADILNEKIREELGRI 660
Qy 1000 EQHREKELEVQKQLEALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 1059
Db 661 EQHREKELEVQKQLEALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKKEESLTKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKKEESLTKRASQYSGQLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 1179
Db 781 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 840
Qy 1180 NVDVSTIYNNEVLHQPILSEAQKSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 841 NVDVSTIYNNEVLHQPILSEAQKSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 900
Qy 1240 KEAEHMYQNEQDNVANKHTEQOESLDQKLQLOSKNMWLQOOLVHAHKADNKSKITIDIH 1299
Db 901 KEAEHMYQNEQDNVANKHTEQOESLDQKLQLOSKNMWLQOOLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKXQHLLKKEKNEEIFNYYNNHLKNRIQYKEKEAETENS 1341
Db 961 FLERKXQHLLKKEKNEEIFNYYNNHLKNRIQYKEKEAETENS 1002
```

RESULT 13

```
US-10-076-622-475
; Sequence 475, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; CURRENT FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 627
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-076-622-475
```

```
Query Match 74.1%; Score 5130; DB 14; Length 1002;
Best Local Similarity 99.4%; Pred. No. 1.6e-25;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 340 MSPAKETSEKFTWAAGRPRIKIAWEKKEPTPVKTGCVARVTSNKTAVLEKGRSKMTACPTK 399
Db 1 MSPAKETSEKFTWAAGRPRIKIAWEKKEPTPVKTGCVARVTSNKTAVLEKGRSKMTACPTK 60
Qy 400 ESSTKASANDQRFPSSEKQEEDEEYSCDSRSLFESSAKIQVCIPESIIQKVMIEINREVEE 459
Db 61 ESSTKASANDQRFPSSEKQEEDEEYSCDSRSLFESSAKIQVCIPESIIQKVMIEINREVEE 120
Qy 460 PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPESSKQKQYENSWSSESICETV 519
Db 121 PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPESSKQKQYENSWSSESICETV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 240
Qy 580 PKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 639
Db 241 PKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 360
Qy 700 EKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 759
Db 361 EKPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 420
Qy 760 KDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 819
Db 421 KDVCLPKATHQKEIDKINGKLEESPNKDGGLKATCGMKVSIPTKALELMDMOTFKABP 480
Qy 820 KPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 879
Db 481 KPSAFEPATEMOKSVPNKALELKNQOTLRADQELIPSESKQKQYENSWSSESICETVS 540
Qy 880 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
Qy 940 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKXKLSEAKEIKSOLQENQKQWQELCSVRLTNQEEKRRNADILNEKIREELGRI 660
Qy 1000 EQHREKELEVQKQLEALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 1059
Db 661 EQHREKELEVQKQLEALRIQDIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKKEESLTKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYFEDIKILKEKNAELQMTLKKEESLTKRASQYSGQLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 1179
Db 781 TMLTSKLKEKQDKEILEAEIESHHPRLASAVQDHDQIVTSRKSPAFHAGDACLQKRM 840
Qy 1180 NVDVSTIYNNEVLHQPILSEAQKSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 841 NVDVSTIYNNEVLHQPILSEAQKSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 900
```


Db 841 NVDVSTIYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAEHMYQNEQDNNKHTQEESLDQKLFQLOSKMMLQOOLVHAHKADNKSKITIDIH 1299
Db 901 KEAEHMYQNEQDNNKHTQEESLDQKLFQLOSKMMLQOOLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKMQHLLKEKNEEIFNYYNNHLKNRIYQYEKEAETENS 1341
Db 961 FLERKMQHLLKEKNEEIFNYYNNHLKNRIYQYEKEAETENS 1002

RESULT 14
US-10-124-805-475
; Sequence 475, Application US/10124805
; Publication No. US20030166022A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C12
; CURRENT APPLICATION NUMBER: US/10/124,805
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 310, 429, 522
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-124-805-475

Query Match 74.1%; Score 5130; DB 14; Length 1002;
Best Local Similarity 99.4%; Pred. No. 1.6e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTPVKTGCVARVTSNKTIVLEKGRSKMTACPTK 399
Db 1 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTPVKTGCVARVTSNKTIVLEKGRSKMTACPTK 60
Qy 400 ESTKASANDQRPFSKOEDEEYSCDSRSIFESSAKIOVCIPESIQKWEINREVEE 459
Db 61 ESTKASANDQRPFSKOEDEEYSCDSRSIFESSAKIOVCIPESIQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 519
Db 121 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 180
Qy 520 SQKDVCLPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 579
Db 181 SQKDVCLPKATHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 240
Qy 580 PKGPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 639
Db 241 PKGPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 300
Qy 640 QKDVCLPKAAHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 699
Db 301 QKDVCLPKAAHQKIDKINGKLESPNKGKLLKATCGMKVSIPTKALELMDMTFKAEPP 360
Qy 700 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 759
Db 361 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 420
Qy 760 KDVCLPKATHQKIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMTFKAEPP 819
Db 421 KDVCLPKATHQKIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMTFKAEPP 480
Qy 820 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 879

Db 481 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADPMFPPEKQKDYEENSWSLCEV 540
Qy 880 DVCVPKATHQKIDKINGKLESDSTLSKILDTVHSCERARELOKDHCEQRTGMEQMKK 939
Db 541 DVCVPKATHQKIDKINGKLESDSTLSKILDTVHSCERARELOKDHCEQRTGMEQMKK 600
Qy 940 FCVLKKLSEAKETKSQLENQKVKWEQELCSVRLTLNQEEBKRNADILNEKIREELGRI 999
Db 601 FCVLKKLSEAKETKSQLENQKVKWEQELCSVRLTLNQEEBKRNADILNEKIREELGRI 660
Qy 1000 EEQHRKELEVQKQLEQALRIQDIELKSVEENLQVSHHNNYLLHENCWMLKKEIAMLK 1059
Db 661 EEQHRKELEVQKQLEQALRIQDIELKSVEENLQVSHHNNYLLHENCWMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGOLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGOLKVLIAEN 780
Qy 1120 TMLTSKLKQKQKKEILEAEIESHHPRLASAVQDHQIVTSRKSQEPAFHAGDACLQRM 1179
Db 781 TMLTSKLKQKQKKEILEAEIESHHPRLASAVQDHQIVTSRKSQEPAFHAGDACLQRM 840
Qy 1180 NVDVSTIYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 1239
Db 841 NVDVSTIYNNEVLHQLPSEARQSKSLKINLVAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAEHMYQNEQDNNKHTQEESLDQKLFQLOSKMMLQOOLVHAHKADNKSKITIDIH 1299
Db 901 KEAEHMYQNEQDNNKHTQEESLDQKLFQLOSKMMLQOOLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKMQHLLKEKNEEIFNYYNNHLKNRIYQYEKEAETENS 1341
Db 961 FLERKMQHLLKEKNEEIFNYYNNHLKNRIYQYEKEAETENS 1002

RESULT 15
US-09-834-759-493
; Sequence 493, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493

Query Match 74.1%; Score 5130; DB 9; Length 1095;
Best Local Similarity 99.4%; Pred. No. 1.8e-257;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTPVKTGCVARVTSNKTIVLEKGRSKMTACPTK 399
Db 94 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTPVKTGCVARVTSNKTIVLEKGRSKMTACPTK 153

QY 400 ESSTKASANDQRPSESQEBDEEYSCDSRSLPSSAKIQVCIPESIIYQKWEINREVEE 459
DB |||||
DB 154 ESSTKASANDQRPSESQEBDEEYSCDSRSLPSSAKIQVCIPESIIYQKWEINREVEE 213
QY 460 PPKPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 519
DB |||||
DB 214 PPKPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 273
QY 520 SQKDVCPLKATHQKEIDKINGKLEBSPNKDGLLKATCGMKVSIPTKALELXDMQTFKAEP 579
DB |||||
DB 274 SQKDVCPLKATHQKEIDKINGKLEBSPNKDGLLKATCGMKVSIPTKALELXDMQTFKAEP 333
QY 580 PGKPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 639
DB |||||
DB 334 PGKPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 393
QY 640 QKDVCPLKAAHQKEIDKINGKLEBSPNKDGLLKATCGMKVSIPTKALELXDMQTFKAEP 699
DB |||||
DB 394 QKDVCPLKAAHQKEIDKINGKLEBSPNKDGLLKATCGMKVSIPTKALELXDMQTFKAEP 453
QY 700 EKPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 759
DB |||||
DB 454 EKPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 513
QY 760 KDVCLPKATHQKEIDKINGKLEBSPNDGFLKAPCRMVSIPTKALELXDMQTFKAEP 819
DB |||||
DB 514 KDVCLPKATHQKEIDKINGKLEBSPNDGFLKAPCRMVSIPTKALELXDMQTFKAEP 573
QY 820 KPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 879
DB |||||
DB 574 KPSAFKPALEMONSVPNKAFELKNEOTLRADMPFPESKQDYEESSWDSLSLCTV 633
QY 880 DVCVPKATHQKEMDKISGLKEDSTSLKILDTVHSCERARELOKHCEQRTGKMEQMKK 939
DB |||||
DB 634 DVCVPKATHQKEMDKISGLKEDSTSLKILDTVHSCERARELOKHCEQRTGKMEQMKK 693
QY 940 FCVLKKLSEAKEIKSQLENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
DB |||||
DB 694 FCVLKKLSEAKEIKSQLENQKQWQELCSVRLTLNQEEKRRNADILNEKIREELGRI 753
QY 1000 EEOHRELKQVQLEQALRIQDIELKSVEINLQVSHTHENENYLLHENCMLKKEIAMLK 1059
DB |||||
DB 754 EEOHRELKQVQLEQALRIQDIELKSVEINLQVSHTHENENYLLHENCMLKKEIAMLK 813
QY 1060 LEIATLKHQYQKQENKYPEDIKILKEKNAELQMTLKEESLTKRASQYSGOLKVLIAEN 1119
DB |||||
DB 814 LEIATLKHQYQKQENKYPEDIKILKEKNAELQMTLKEESLTKRASQYSGOLKVLIAEN 873
QY 1120 TMLTSKLKQDKQKEILEAEIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 1179
DB |||||
DB 874 TMLTSKLKQDKQKEILEAEIESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 933
QY 1180 NYDVSTTYNNEVLHOPLEAQRKSLSKINLYAGDALRENTLVSEHAQDQRETQOM 1239
DB |||||
DB 934 NYDVSTTYNNEVLHOPLEAQRKSLSKINLYAGDALRENTLVSEHAQDQRETQOM 993
QY 1240 KEAEHMYQNEQDQVNNKHTQESLQKLFQLSKNMWLQOOLVHAHKADNKSKITIDIH 1299
DB |||||
DB 994 KEAEHMYQNEQDQVNNKHTQESLQKLFQLSKNMWLQOOLVHAHKADNKSKITIDIH 1053
QY 1300 FLERKQHHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
DB |||||
DB 1054 FLERKQHHLLKKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1095

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:57:06 ; Search time 36.5216 Seconds
(without alignments)
2740.961 Million cell updates/sec

Title: US-09-602-362E-23

Perfect score: 6920

Sequence: 1 MTRKKTINLQDAQRKA.....NHLKNRIYQYEKAEKATENS 1341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5130	74.1	1002	4	US-09-620-405B-475
2	5130	74.1	1002	4	US-09-604-287A-475
3	5130	74.1	1002	4	US-09-834-759-475
4	5130	74.1	1002	4	US-09-590-751A-475
5	5130	74.1	1002	4	US-09-551-621-475
6	5130	74.1	1095	4	US-09-620-405B-493
7	5130	74.1	1095	4	US-09-834-759-493
8	3252.5	47.0	650	4	US-09-620-405B-469
9	3252.5	47.0	650	4	US-09-433-826B-469
10	3252.5	47.0	650	4	US-09-604-287A-469
11	3252.5	47.0	650	4	US-09-834-759-469
12	3252.5	47.0	650	4	US-09-590-751A-469
13	3252.5	47.0	650	4	US-09-551-621-469
14	3252.5	47.0	743	4	US-09-620-405B-494
15	3252.5	47.0	743	4	US-09-834-759-494
16	2589	37.4	512	4	US-09-451-739H-16
17	2270	32.8	466	4	US-09-620-405B-472
18	2270	32.8	466	4	US-09-433-826B-472
19	2270	32.8	466	4	US-09-604-287A-472
20	2270	32.8	466	4	US-09-834-759-472
21	2270	32.8	466	4	US-09-590-751A-472
22	2270	32.8	466	4	US-09-551-621-472
23	2256	32.6	445	4	US-09-620-405B-473
24	2256	32.6	445	4	US-09-433-826B-473
25	2256	32.6	445	4	US-09-604-287A-473
26	2256	32.6	445	4	US-09-834-759-473
27	2256	32.6	445	4	US-09-590-751A-473

28	2256	32.6	445	4	US-09-551-621-473	Sequence 473, App
29	2212	32.0	432	4	US-09-389-681-181	Sequence 181, App
30	2212	32.0	432	4	US-09-620-405B-181	Sequence 181, App
31	2212	32.0	432	4	US-09-339-338-181	Sequence 181, App
32	2212	32.0	432	4	US-09-433-826B-181	Sequence 181, App
33	2212	32.0	432	4	US-09-604-287A-181	Sequence 181, App
34	2212	32.0	432	4	US-09-285-480-181	Sequence 181, App
35	2212	32.0	432	4	US-09-834-759-181	Sequence 181, App
36	2212	32.0	432	4	US-09-590-751A-181	Sequence 181, App
37	2212	32.0	432	4	US-09-551-621-181	Sequence 181, App
38	1583.5	22.9	410	4	US-09-620-405B-495	Sequence 495, App
39	1583.5	22.9	410	4	US-09-834-759-495	Sequence 495, App
40	1576	22.8	317	4	US-09-389-681-176	Sequence 176, App
41	1576	22.8	317	4	US-09-620-405B-176	Sequence 176, App
42	1576	22.8	317	4	US-09-339-338-176	Sequence 176, App
43	1576	22.8	317	4	US-09-433-826B-176	Sequence 176, App
44	1576	22.8	317	4	US-09-604-287A-176	Sequence 176, App
45	1576	22.8	317	4	US-09-285-480-176	Sequence 176, App

ALIGNMENTS

RESULT 1

US-09-620-405B-475
; Sequence 475, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-475

Query Match		74.1%;	Score 5130;	DB 4;	Length 1002;
Best Local Similarity		99.4%;	Pred. No. 0;		
Matches		996;	Conservative	2;	Mismatches 4; Indels 0; Gaps 0;
Qy	340	MSPAKETSEKFTAAKGRPRKIAWEKETPVKTGCVARVTGNKTKVLEKGRSKMIACPTK	399		
Db	1	MSPAKETSEKFTAAKGRPRKIAWEKETPVKTGCVARVTGNKTKVLEKGRSKMIACPTK	60		
Qy	400	ESSTKASANDQRPSPSKQDEEYSCDSRLFPSSAKIQVCIPESIQKVMENREVVEE	459		
Db	61	ESSTKASANDQRPSPSKQDEEYSCDSRLFPSSAKIQVCIPESIQKVMENREVVEE	120		
Qy	460	PPKPSAPKPAIENQNSVPNKAFELKNEQTLRADPMFPSPSKQDEYNSWDSLSLCTV	519		
Db	121	PPKPSAPKPAIENQNSVPNKAFELKNEQTLRADPMFPSPSKQDEYNSWDSLSLCTV	180		
Qy	520	SQKDVCLPKATHQEKIDKINGKLBESPNKDGLLKATCGMKVSIPTKALELXDMQTFKAEP	579		
Db	181	SQKDVCLPKATHQEKIDKINGKLBESPNKDGLLKATCGMKVSIPTKALELXDMQTFKAEP	240		
Qy	580	PGKPSAFEPATEMOKSVNKALELKNQETWRADEILPSPSKQDEYNSWDSLSLCTV	639		

```

Db 241 PGKPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVS 300
Qy 640 QKDVCLPKAAHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 699
Db 301 QKDVCLPKAAHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 360
Qy 700 EKPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 759
Db 361 EKPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 420
Qy 760 KDVCLPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 819
Db 421 KDVCLPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 480
Qy 820 KPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 879
Db 481 KPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 540
Qy 880 DVCVPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 939
Db 541 DVCVPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 600
Qy 940 FCVLKXKLESEAKEIKSQLENOKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKXKLESEAKEIKSQLENOKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 BEQHRKELEVKQOQLEALRIODIELKSVESSLNQVSHTHENYLLHNCMLKKEIAMLK 1059
Db 661 BEQHRKELEVKQOQLEALRIODIELKSVESSLNQVSHTHENYLLHNCMLKKEIAMLK 720
Qy 1060 LEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQRM 1179
Db 781 TMLTSKLKEKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQRM 840
Qy 1180 NYDVSSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 1239
Db 841 NYDVSSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 900
Qy 1240 KEAHEMYQNEQDNVKNKHTEQESLQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 1299
Db 901 KEAHEMYQNEQDNVKNKHTEQESLQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 960

```

```

RESULT 2
US-09-604-287A-475
; Sequence 475. Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hessler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604, 287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien

```

```

; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-475

Query Match 74.1%; Score 5130; DB 4; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEKTPVKTGCVARVTSNKTVKLEGRSOMIACTPK 399
Db 1 MSPAKETSEKFTWAAKGRPRKIAWEKKEKTPVKTGCVARVTSNKTVKLEGRSOMIACTPK 60
Qy 400 ESSTKASANDORPPSPESKQDEEYSCDSRSLPESAKIQVCIPESIIYOKVMEINREVEE 459
Db 61 ESSTKASANDORPPSPESKQDEEYSCDSRSLPESAKIQVCIPESIIYOKVMEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVNPNAFELKNEQTLRADPMPFPSPESKQDYESSWDSSELCTV 519
Db 121 PPKPSAFKPAIEMQNSVNPNAFELKNEQTLRADPMPFPSPESKQDYESSWDSSELCTV 180
Qy 520 SQKDVCLPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 579
Db 181 SQKDVCLPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 240
Qy 580 PGKPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVS 639
Db 241 PGKPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVS 300
Qy 640 QKDVCLPKAAHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 699
Db 301 QKDVCLPKAAHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 360
Qy 700 EKPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 759
Db 361 EKPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 420
Qy 760 KDVCLPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 819
Db 421 KDVCLPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 480
Qy 820 KPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 879
Db 481 KPSAFEPATEMOKSVNKALELNQETLRADILPSESKQDYESSWDSSELCTVSQ 540
Qy 880 DVCVPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 939
Db 541 DVCVPKATHQEKIDKINGKLESPVKGDLKANKCMKVSIPTKALELMDMOTFFKAEP 600
Qy 940 FCVLKXKLESEAKEIKSQLENOKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 999
Db 601 FCVLKXKLESEAKEIKSQLENOKVWEQELCSVRLTLNQEEKRRNADILNEKIREELGRI 660
Qy 1000 BEQHRKELEVKQOQLEALRIODIELKSVESSLNQVSHTHENYLLHNCMLKKEIAMLK 1059
Db 661 BEQHRKELEVKQOQLEALRIODIELKSVESSLNQVSHTHENYLLHNCMLKKEIAMLK 720
Qy 1060 LEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHOYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 780
Qy 1120 TMLTSKLKEKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQRM 1179
Db 781 TMLTSKLKEKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQBPAPHIAGDACLQRM 840
Qy 1180 NYDVSSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 1239
Db 841 NYDVSSTIYNNEVLHQPSEAKRSKSLKINLYAGDALRENTLYSEHAQDQRETQCOM 900
Qy 1240 KEAHEMYQNEQDNVKNKHTEQESLQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 1299
Db 901 KEAHEMYQNEQDNVKNKHTEQESLQKLFQOSKNMWLQOQLVHAHKADNKSKITIDIH 960

```

```
Qy 1300 FLERKMQHLLKKEKNEBIFNYYNNHLLKNRIYQYEKEKAETENS 1341
Db 961 FLERKMQHLLKKEKNEBIFNYYNNHLLKNRIYQYEKEKAETENS 1002

RESULT 3
US-09-834-759-475
; Sequence 475, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-834-759-475

Query Match 74.1%; Score 5130; DB 4; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAGPRKIAWEKKEKTPVKTGCVARVTSNKTKEGSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGPRKIAWEKKEKTPVKTGCVARVTSNKTKEGSKMIACPTK 60

Qy 400 ESSTKASANDQRPFSSEKQDEEYSCDSRSLFESSAKIQVICIPESIQKWEINREVEE 459
Db 61 ESSTKASANDQRPFSSEKQDEEYSCDSRSLFESSAKIQVICIPESIQKWEINREVEE 120

Qy 460 PPKPSAPKPAIEMQNSVFNKAFELKNEQTLRADPMPFPSPESKQDYENSWDSLSLCTV 519
Db 121 PPKPSAPKPAIEMQNSVFNKAFELKNEQTLRADPMPFPSPESKQDYENSWDSLSLCTV 180

Qy 520 SOKDVCLPKATHQKEIDKINGKLESPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 579
Db 181 SOKDVCLPKATHQKEIDKINGKLESPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 240

Qy 580 PGKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVS 639
Db 241 PGKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVS 300

Qy 640 QKDVCLPKAAHQEKIDKINGKLESPVKGGLLKANCCKMKVSIPTKALELMDQMTFKAEP 699
Db 301 QKDVCLPKAAHQEKIDKINGKLESPVKGGLLKANCCKMKVSIPTKALELMDQMTFKAEP 360

Qy 700 EKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVSQ 759
Db 361 EKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVSQ 420

Qy 760 KDVCPLKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEP 819
Db 421 KDVCPLKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEP 480

Qy 820 KPSAFEPATEMOKSVNPKALELKNQETLRADQMPFSPESKQDYENSWDSLSLCTVSQ 879
Db 400 KPSAFEPATEMOKSVNPKALELKNQETLRADQMPFSPESKQDYENSWDSLSLCTVSQ 120

Db 481 KPSAFEPATEMOKSVNPKALELKNQETLRADQMPFSPESKQDYENSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLEDSTSLSKILTDTVHSCERABELQKDHCEQRTGKMEQMKKK 939
Db 541 DVCVPKATHQKEMDKISGKLEDSTSLSKILTDTVHSCERABELQKDHCEQRTGKMEQMKKK 600

Qy 940 FCVLKKKLSEAKEIKSQLENQKVKWEQELCSVRLTLNOEERKRRNADILNEKIRELGRI 999
Db 601 FCVLKKKLSEAKEIKSQLENQKVKWEQELCSVRLTLNOEERKRRNADILNEKIRELGRI 660

Qy 1000 BEQHRKELEVKQQLQEQALRIQDIELKSVESNLNVSHTHENYLLHNCMLKKEIAMLK 1059
Db 661 BEQHRKELEVKQQLQEQALRIQDIELKSVESNLNVSHTHENYLLHNCMLKKEIAMLK 720

Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 780

Qy 1120 TMLTSKLKEKQDKKEILEABIESHPRLASAVQDHDQIVTSRKSQEPAPHIAAGDACLQRM 1179
Db 781 TMLTSKLKEKQDKKEILEABIESHPRLASAVQDHDQIVTSRKSQEPAPHIAAGDACLQRM 840

Qy 1180 NVDVSSITYNNEVLHOPLESEAQRKSKILNLYAGDALRENTLVSEHAQDQRETCOM 1239
Db 841 NVDVSSITYNNEVLHOPLESEAQRKSKILNLYAGDALRENTLVSEHAQDQRETCOM 900

Qy 1240 KEASHMYQNEQDNVKNKITEQOESLDQKLFQSQKNMQLQQLVHAHKKADNKSKITIDH 1299
Db 901 KEASHMYQNEQDNVKNKITEQOESLDQKLFQSQKNMQLQQLVHAHKKADNKSKITIDH 960

Qy 1300 FLERKMQHLLKKEKNEBIFNYYNNHLLKNRIYQYEKEKAETENS 1341
Db 961 FLERKMQHLLKKEKNEBIFNYYNNHLLKNRIYQYEKEKAETENS 1002

RESULT 4
US-09-590-751A-475
; Sequence 475, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-590-751A-475

Query Match 74.1%; Score 5130; DB 4; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAGPRKIAWEKKEKTPVKTGCVARVTSNKTKEGSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGPRKIAWEKKEKTPVKTGCVARVTSNKTKEGSKMIACPTK 60

Qy 400 ESSTKASANDQRPFSSEKQDEEYSCDSRSLFESSAKIQVICIPESIQKWEINREVEE 459
Db 61 ESSTKASANDQRPFSSEKQDEEYSCDSRSLFESSAKIQVICIPESIQKWEINREVEE 120

Qy 460 PPKPSAPKPAIEMQNSVFNKAFELKNEQTLRADPMPFPSPESKQDYENSWDSLSLCTV 519
Db 121 PPKPSAPKPAIEMQNSVFNKAFELKNEQTLRADPMPFPSPESKQDYENSWDSLSLCTV 180

Qy 520 SOKDVCLPKATHQKEIDKINGKLESPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 579
Db 181 SOKDVCLPKATHQKEIDKINGKLESPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 240

Qy 580 PGKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVS 639
Db 241 PGKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVS 300

Qy 640 QKDVCLPKAAHQEKIDKINGKLESPVKGGLLKANCCKMKVSIPTKALELMDQMTFKAEP 699
Db 301 QKDVCLPKAAHQEKIDKINGKLESPVKGGLLKANCCKMKVSIPTKALELMDQMTFKAEP 360

Qy 700 EKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVSQ 759
Db 361 EKPSAFEPATEMOKSVNPKALELKNQETWRADELPSSEKQDYENSWDSLSLCTVSQ 420

Qy 760 KDVCPLKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEP 819
Db 421 KDVCPLKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEP 480

Qy 820 KPSAFEPATEMOKSVNPKALELKNQETLRADQMPFSPESKQDYENSWDSLSLCTVSQ 879
Db 400 KPSAFEPATEMOKSVNPKALELKNQETLRADQMPFSPESKQDYENSWDSLSLCTVSQ 120
```

QY 460 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 519
DB 121 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 180
QY 520 SOKVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 579
DB 181 SOKVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 240
QY 580 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 639
DB 241 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 300
QY 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 699
DB 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 360
QY 700 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 759
DB 361 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 420
QY 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 819
DB 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 480
QY 820 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 879
DB 481 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 540
QY 880 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 939
DB 541 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 600
QY 940 FCVLKPKLSEAKETKSOLENQKVEOBLCSVRILTNOEERKRRNADILNEKIREELGRI 999
DB 601 FCVLKPKLSEAKETKSOLENQKVEOBLCSVRILTNOEERKRRNADILNEKIREELGRI 660
QY 1000 EEOHRELKVEKQLEQALRIODIELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLK 1059
DB 661 EEOHRELKVEKQLEQALRIODIELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLK 720
QY 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 1119
DB 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 780
QY 1120 TMLTSKLKEKQDKKEILEAEIESHHPRLASAVODHDQIVTSRKSQEPAFHAGDACLQRM 1179
DB 781 TMLTSKLKEKQDKKEILEAEIESHHPRLASAVODHDQIVTSRKSQEPAFHAGDACLQRM 840
QY 1180 NVDVSTTYNNEVLHQPILSEAKQKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 1239
DB 841 NVDVSTTYNNEVLHQPILSEAKQKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 900
QY 1240 KEAEMYQNEQDNVKNKTEQESLQKLFOLOKNNWLQOQLVHAHKADNKSKITIDIH 1299
DB 901 KEAEMYQNEQDNVKNKTEQESLQKLFOLOKNNWLQOQLVHAHKADNKSKITIDIH 960
QY 1300 FLERQVQHLLKKEKEEIEFNYNHLLKNRIYQYEKEKAETENS 1341
DB 961 FLERQVQHLLKKEKEEIEFNYNHLLKNRIYQYEKEKAETENS 1002

RESULT 5

US-09-551-621-475
; Sequence 475, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqi, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470CS
; CURRENT APPLICATION NUMBER: US/09/551.621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 475
; LENGTH: 1002
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1002)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-475

Query Match 74.1%; Score 5130; DB 4; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 340 MSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTGCVARVTSNKTKVLEKGRSKMIACPTK 399
DB 1 MSPAKETSEKFTWAAKGRPRKIAWEKKETPVKTGCVARVTSNKTKVLEKGRSKMIACPTK 60
QY 400 ESSTKASANDQRPSESQBEDEYSCDSRLPSSAKIQVCIPESIQKWEINREVVEE 459
DB 61 ESSTKASANDQRPSESQBEDEYSCDSRLPSSAKIQVCIPESIQKWEINREVVEE 120
QY 460 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 519
DB 121 PPKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 180
QY 520 SOKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 579
DB 181 SOKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 240
QY 580 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 639
DB 241 PKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 300
QY 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 699
DB 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 360
QY 700 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 759
DB 361 EKPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 420
QY 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 819
DB 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 480
QY 820 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 879
DB 481 KPSAFKPAIEMQNSVFNKAFELKNEQTLRADMPFPESKQKDYENSWSLSLCTV 540
QY 880 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 939
DB 541 DVCVPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 600
QY 940 FCVLKPKLSEAKETKSOLENQKVEOBLCSVRILTNOEERKRRNADILNEKIREELGRI 999
DB 601 FCVLKPKLSEAKETKSOLENQKVEOBLCSVRILTNOEERKRRNADILNEKIREELGRI 660
QY 1000 EEOHRELKVEKQLEQALRIODIELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLK 1059
DB 661 EEOHRELKVEKQLEQALRIODIELKSVESNLNOVSHTHENYLLHNCMLKKEIAMLK 720
QY 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 1119
DB 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSQLKVLIAEN 780
QY 1120 TMLTSKLKEKQDKKEILEAEIESHHPRLASAVODHDQIVTSRKSQEPAFHAGDACLQRM 1179

```
Db 781 TMLTSKLEKODKELEAEIESHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 840
Qy 1180 NVDVSTIYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 1239
Db 841 NVDVSTIYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAEMHYQNEQDNVKNKTEQESLQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 1299
Db 901 KEAEMHYQNEQDNVKNKTEQESLQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 960
Qy 1300 FLERKQMHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 961 FLERKQMHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1002

RESULT 6
US-09-620-405B-493
; Sequence 493, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-493

Query Match 74.1%; Score 5130; DB 4; Length 1095;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEKTPVKTGCVARVTSNKTIVLEKGRSKMIACPTK 399
Db 94 MSPAKETSEKFTWAAKGRPRKIAWEKKEKTPVKTGCVARVTSNKTIVLEKGRSKMIACPTK 153
Qy 400 ESSTKASANDORFPSESQDEEYSCDSRSLFESSAKIOVCIPESIIYOKWEINREVEE 459
Db 154 ESSTKASANDORFPSESQDEEYSCDSRSLFESSAKIOVCIPESIIYOKWEINREVEE 213
Qy 460 PPKPSAFKPAEMQNSVPNKAFELKNEQTLRADPMFPPEKQKDYENSWSLCLCTV 519
Db 214 PPKPSAFKPAEMQNSVPNKAFELKNEQTLRADPMFPPEKQKDYENSWSLCLCTV 273
Qy 520 SOKVCLPKATHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDMTFFKAEP 579
Db 274 SOKVCLPKATHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDMTFFKAEP 333
Qy 580 PKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWSLCLCTVS 639
Db 334 PKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWSLCLCTVS 393
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCGRMKVSIPTKALELMDMTFFKAEP 699
Db 394 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCGRMKVSIPTKALELMDMTFFKAEP 453
Qy 700 EKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWSLCLCTVSQ 759
```

```
Db 454 EKPSAFEPATMOKSVPNKALELNEQTLRADEILPSESQKDYENSWSLCLCTVSQ 513
Qy 760 KDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFFKAEP 819
Db 514 KDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFFKAEP 573
Qy 820 KPSAFEPATMOKSVPNKALELNEQTLRADQMPFSESQKDYENSWSLCLCTVSQ 879
Db 574 KPSAFEPATMOKSVPNKALELNEQTLRADQMPFSESQKDYENSWSLCLCTVSQ 633
Qy 880 DVCVPKATHQEMDKISGKLESDSTSLKILDTVHSCERARELOKDHQRTGKMEQMKK 939
Db 634 DVCVPKATHQEMDKISGKLESDSTSLKILDTVHSCERARELOKDHQRTGKMEQMKK 693
Qy 940 FCVLKCKLSEAKEIKSOLENQKVAEQELCSVRLTLNQEERKRNADTLNEKIREELGRI 999
Db 694 FCVLKCKLSEAKEIKSOLENQKVAEQELCSVRLTLNQEERKRNADTLNEKIREELGRI 753
Qy 1000 EEOHKELEVKQOLEQALRIQDIIEKSVESNLNQVSHTHENYLLHENCMLKKEIAMLK 1059
Db 754 EEOHKELEVKQOLEQALRIQDIIEKSVESNLNQVSHTHENYLLHENCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPFEDIKILKEKVAELQMTLKLKEESITKRASQYSGOLKVLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPFEDIKILKEKVAELQMTLKLKEESITKRASQYSGOLKVLIAEN 873
Qy 1120 TMLTSKLEKODKEILEAEIESHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 1179
Db 874 TMLTSKLEKODKEILEAEIESHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRM 933
Qy 1180 NVDVSTIYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 1239
Db 934 NVDVSTIYNNEVLHQPSEARQSKSLKINLYAGDALRENTLVSEHAQRDQRETQCM 993
Qy 1240 KEAEMHYQNEQDNVKNKTEQESLQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 1299
Db 994 KEAEMHYQNEQDNVKNKTEQESLQKLFQLOSKNMWLOQQLVHAHKADNKSKITIDH 1053
Qy 1300 FLERKQMHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1341
Db 1054 FLERKQMHLLKEKNEEIFNYYNNHLKNRIYQYEKEKAETENS 1095
```

```
RESULT 7
US-09-834-759-493
; Sequence 493, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: variant
; LOCATION: (1)...(1095)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-834-759-493
```

Query Match 74.1%; Score 5130; DB 4; Length 1095;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 99%; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 340 MSPAKETSEKFTWAAGRPRIAKWKETPVKTCVARVTSNKTIVLEKGRSKMIACPTK 399
DB 94 MSPAKETSEKFTWAAGRPRIAKWKETPVKTCVARVTSNKTIVLEKGRSKMIACPTK 153

QY 400 ESSTKASANDORFPSEKQDEEYSCDSRSIFESSAKIOVCIPESIQKWEINREVEE 459
DB 154 ESSTKASANDORFPSEKQDEEYSCDSRSIFESSAKIOVCIPESIQKWEINREVEE 213

QY 460 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 519
DB 214 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 273

QY 520 SKQDVCLPKATHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABP 579
DB 274 SKQDVCLPKATHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABP 333

QY 580 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 639
DB 334 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 393

QY 640 QKDVCPLKAAHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABP 699
DB 394 QKDVCPLKAAHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABP 453

QY 700 EKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 759
DB 454 EKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 513

QY 760 KDVCPLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 819
DB 514 KDVCPLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 573

QY 820 KPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 879
DB 574 KPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 633

QY 880 DVCVPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 939
DB 634 DVCVPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 693

QY 940 FCVLAKKLSKAEIKSQLENOQKWEQELCSVRLTLNOEERKRRNADILNEKIREELGRI 999
DB 694 FCVLAKKLSKAEIKSQLENOQKWEQELCSVRLTLNOEERKRRNADILNEKIREELGRI 753

QY 1000 EBOHRKELEVKKQLEALRIQDIELKSVEENLNQVSHTHENENYLLHENCMLKKEIAMLK 1059
DB 754 EBOHRKELEVKKQLEALRIQDIELKSVEENLNQVSHTHENENYLLHENCMLKKEIAMLK 813

QY 1060 LEIATLKHOYQKKNYFEDIKILKEKNAELQMTLKEESLTKRASQVSGQLKVLIAEN 1119
DB 814 LEIATLKHOYQKKNYFEDIKILKEKNAELQMTLKEESLTKRASQVSGQLKVLIAEN 873

QY 1120 TMLTSLKKEQDKEILEAESHHPRLASAVODHDOI VTSRKSQSPAFIAGDACLQRM 1179
DB 874 TMLTSLKKEQDKEILEAESHHPRLASAVODHDOI VTSRKSQSPAFIAGDACLQRM 933

QY 1180 NVDSSTIYNNEVLHQPLSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETOQM 1239
DB 934 NVDSSTIYNNEVLHQPLSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETOQM 993

QY 1240 KEAEMHYQNEQNVNKHTEQBSLQKLFQLOSKNWLQOQLVHAHKADNKSKITIDH 1299
DB 994 KEAEMHYQNEQNVNKHTEQBSLQKLFQLOSKNWLQOQLVHAHKADNKSKITIDH 1053

QY 1300 FLERKMQHLLKKEKEEFNYYNNHLKNRIYQYEKEKAETENS 1341
DB 1054 FLERKMQHLLKKEKEEFNYYNNHLKNRIYQYEKEKAETENS 1095

RESULT 8
US-09-620-405B-469
; Sequence 469, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620.405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-620-405B-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 340 MSPAKETSEKFTWAAGRPRIAKWKETPVKTCVARVTSNKTIVLEKGRSKMIACPTK 399
DB 1 MSPAKETSEKFTWAAGRPRIAKWKETPVKTCVARVTSNKTIVLEKGRSKMIACPTK 60

QY 400 ESSTKASANDORFPSEKQDEEYSCDSRSIFESSAKIOVCIPESIQKWEINREVEE 459
DB 61 ESSTKASANDORFPSEKQDEEYSCDSRSIFESSAKIOVCIPESIQKWEINREVEE 120

QY 460 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 519
DB 121 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 180

QY 520 SKQDVCLPKATHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABP 579
DB 181 SKQDVCLPKATHQKEIDKINGKLESPNKDGLLKATCGMKVSIPTKALELMDQTFKABP 240

QY 580 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 639
DB 241 PKKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 300

QY 640 QKDVCPLKAAHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 699
DB 301 QKDVCPLKAAHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 360

QY 700 EKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 759
DB 361 EKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 420

QY 760 KDVCPLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 819
DB 421 KDVCPLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 480

QY 820 KPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 879
DB 481 KPSAFPAIEMQNSVPNKAFELKNEQTLRADPMFPPEKQDYENSWDSSELCETV 540

QY 880 DVCVPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 939
DB 541 DVCVPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDQTFKABP 600

QY 940 FCVLAKKLSKAEIKSQLENOQKWEQELCSVR-LTL 975


```
Db 601 FCVLKXKLSEAKEIKSQLENOQKVKWEQELCSVR-1TL 637
|||||
RESULT 9
US-09-433-826B-469
; Sequence 469, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-433-826B-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 340 MSPAKETSEKFTWAAGRPRIAKWEKKEPTVKTGCVARVTSNKTQVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGRPRIAKWEKKEPTVKTGCVARVTSNKTQVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDORPSESQDEDEYSCDSRLPESAKIQVICIPESIQKWEINREVEE 459
Db 61 ESSTKASANDORPSESQDEDEYSCDSRLPESAKIQVICIPESIQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPMPPEPSKQDYBENSWSLSCETV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPMPPEPSKQDYBENSWSLSCETV 180
Qy 520 SQDVCLPKATHQKEIDKINGKLESPNKGLLKATCGMKVSIPTKALELMDMTFFKAEP 579
Db 181 SQDVCLPKATHQKEIDKINGKLESPNKGLLKATCGMKVSIPTKALELMDMTFFKAEP 240
Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVS 639
Db 241 PGKPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVS 300
Qy 640 QKDVCPLKAAHQEIDKINGKLESPVKGDLKANKCGMKVSIPTKALELMDMTFFKAEP 699
Db 301 QKDVCPLKAAHQEIDKINGKLESPVKGDLKANKCGMKVSIPTKALELMDMTFFKAEP 360
Qy 700 EKPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVSQ 759
Db 361 EKPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVSQ 420
Qy 760 KDVCPLKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFFKAEP 819
Db 421 KDVCPLKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFFKAEP 480
Qy 820 KPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVSQ 879
Db 481 KPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLESDSTSLSKILDTVHSCERARELQKDHCFQRTGKMQMKKK 939
Db 541 DVCVPKATHQKEMDKISGKLESDSTSLSKILDTVHSCERARELQKDHCFQRTGKMQMKKK 600
Qy 940 FCVLKXKLSEAKEIKSQLENOQKVKWEQELCSVR-1TL 975
Db 601 FCVLKXKLSEAKEIKSQLENOQKVKWEQELCSVR-1TL 637

RESULT 10
US-09-604-287A-469
; Sequence 469, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (429)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 340 MSPAKETSEKFTWAAGRPRIAKWEKKEPTVKTGCVARVTSNKTQVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGRPRIAKWEKKEPTVKTGCVARVTSNKTQVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDORPSESQDEDEYSCDSRLPESAKIQVICIPESIQKWEINREVEE 459
Db 61 ESSTKASANDORPSESQDEDEYSCDSRLPESAKIQVICIPESIQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPMPPEPSKQDYBENSWSLSCETV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPMPPEPSKQDYBENSWSLSCETV 180
Qy 520 SQDVCLPKATHQKEIDKINGKLESPNKGLLKATCGMKVSIPTKALELMDMTFFKAEP 579
Db 181 SQDVCLPKATHQKEIDKINGKLESPNKGLLKATCGMKVSIPTKALELMDMTFFKAEP 240
Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVS 639
Db 241 PGKPSAFEPATEMOKSVPNKALELKNQETWRADILPSESQKQDYBENSWSLSCETVS 300
Qy 640 QKDVCPLKAAHQEIDKINGKLESPVKGDLKANKCGMKVSIPTKALELMDMTFFKAEP 699
Db 301 QKDVCPLKAAHQEIDKINGKLESPVKGDLKANKCGMKVSIPTKALELMDMTFFKAEP 360
```

Qy 700 EKPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 759
Db 361 EKPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 819
Db 421 KDVCPLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTGKMEQMKK 600
Qy 940 FCVLKKLSEAKEIKSOLEKQKWEQBLCSVR-LTL 975
Db 601 FCVLKKLSEAKEIKSOLEKQKWEQBLCSVRFLTL 637

RESULT 11

US-09-834-759-469
; Sequence 469, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yudi
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (429)
; OTHER INFORMATION: Xaa = Any Amino Acid
; NAME/KEY: unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-834-759-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 340 MSPAKETSEKFTWAAGRPRIAEKKEPTVKTGCVARVTNKTVKLEGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGRPRIAEKKEPTVKTGCVARVTNKTVKLEGRSKMIACPTK 60
Qy 400 ESSTKASANDQRPFSKQEEDEEYSCDSRLFPSSAKIQVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDQRPFSKQEEDEEYSCDSRLFPSSAKIQVCIPESIIYQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTV 180
Qy 520 SQKDVCPLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMOTFKAEPP 579

Db 181 SQKDVCPLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMOTFKAEPP 240
Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVS 639
Db 241 PGKPSAFEPATEMOKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVS 300
Qy 640 QKDVCPLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMOTFKAEPP 699
Db 301 QKDVCPLPKATHQKEIDKINGKLESPNKGILLKATCGMKVSIPTKALELMDMOTFKAEPP 360
Qy 700 EKPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 759
Db 361 EKPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 420
Qy 760 KDVCPLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 819
Db 421 KDVCPLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 480
Qy 820 KPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLEDSTLSKILDTVHSCERARELQKHCEQRTGKMEQMKK 600
Qy 940 FCVLKKLSEAKEIKSOLEKQKWEQBLCSVR-LTL 975
Db 601 FCVLKKLSEAKEIKSOLEKQKWEQBLCSVRFLTL 637

RESULT 12

US-09-590-751A-469
; Sequence 469, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(650)
; OTHER INFORMATION: Xaa = Any Amino Acid

US-09-590-751A-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 340 MSPAKETSEKFTWAAGRPRIAEKKEPTVKTGCVARVTNKTVKLEGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGRPRIAEKKEPTVKTGCVARVTNKTVKLEGRSKMIACPTK 60
Qy 400 ESSTKASANDQRPFSKQEEDEEYSCDSRLFPSSAKIQVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDQRPFSKQEEDEEYSCDSRLFPSSAKIQVCIPESIIYQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQKSVPNKALELKNQOTLRADILPSESKQKDYESSWDSLSLCTV 519

Db 121 PPKPSAFKPAIEMQNSVPNKAFLKNEQTLRADPMFPPESSKQKDYESSWDSLSLCTV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 240
Qy 580 PKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVS 639
Db 241 PKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVS 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 360
Qy 700 EKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 759
Db 361 EKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 420
Qy 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 819
Db 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 480
Qy 820 KPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 879
Db 481 KPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKHCHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKHCHCEQRTGKMEQMKK 600
Qy 940 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVR-LTL 975
Db 601 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVRFLT 637

RESULT 13

US-09-551-621-469.
; Sequence 469, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551.621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 469
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (310)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (429)
; NAME/KEY: Xaa = Any Amino Acid<221> unsure
; LOCATION: (522)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-551-621-469

Query Match 47.0%; Score 3252.5; DB 4; Length 650;
Best Local Similarity 98.9%; Pred. No. 7e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 340 MSPAKETSEKFTAAAGPRKIAPKKEPTPVKTGCVARTSNKTKVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTAAAGPRKIAPKKEPTPVKTGCVARTSNKTKVLEKGRSKMIACPTK 60

Qy 400 ESSTKASANDORFPSESKQEEDEEYSCDSRSLFESSAKIQVCIPESIYQKVMIEINREVEE 459
Db 61 ESSTKASANDORFPSESKQEEDEEYSCDSRSLFESSAKIQVCIPESIYQKVMIEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAFLKNEQTLRADPMFPPESSKQKDYESSWDSLSLCTV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAFLKNEQTLRADPMFPPESSKQKDYESSWDSLSLCTV 180
Qy 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 579
Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 240
Qy 580 PKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVS 639
Db 241 PKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVS 300
Qy 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 360
Qy 700 EKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 759
Db 361 EKPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 420
Qy 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 819
Db 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMOTFKAEPP 480
Qy 820 KPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 879
Db 481 KPSAFEPATEMOKSVPNKALELKNEQTLRADEILPSESKQKDYESSWDSLSLCTVSQ 540
Qy 880 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKHCHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELOKHCHCEQRTGKMEQMKK 600
Qy 940 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVR-LTL 975
Db 601 FCVLKXKLSEAKEIKSQLENQKVKWEQELCSVRFLT 637

RESULT 14

US-09-620-405B-494
; Sequence 494, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620.405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(743)
; OTHER INFORMATION: Xaa = Any amino acid
US-09-620-405B-494

Query Match 47.0%; Score 3252.5; DB 4; Length 743;
Best Local Similarity 98.9%; Pred. No. 8.5e-215;
Matches 630; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy	340	MSPAKETSFTWAAGRPRIAMPKETPVKTGCVARVTSNKTUVLEKGRSOWIACTPK	399
Db	94	MSPAKETSFTWAAGRPRIAMPKETPVKTGCVARVTSNKTUVLEKGRSOWIACTPK	153
Qy	400	EGSTKASANDQRPFSQOEDEEYSCDSRSLFESSAKIOVCIPESIIYQKWEINREVEE	459
Db	154	EGSTKASANDQRPFSQOEDEEYSCDSRSLFESSAKIOVCIPESIIYQKWEINREVEE	213
Qy	460	PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMPFPSSKQKDYBENSWDSESICETV	519
Db	214	PPKPSAFPAIEMQNSVPNKAFELKNEQTLRADPMPFPSSKQKDYBENSWDSESICETV	273
Qy	520	SQKDVLCPKATHQEKIDKINGKLBEPSNKDGLLKATCGMKVSIPTKALELMDMOTFKAAP	579
Db	274	SQKDVLCPKATHQEKIDKINGKLBEPSNKDGLLKATCGMKVSIPTKALELMDMOTFKAAP	333
Qy	580	PGKPSAFEPATEMQSVPNKALELKNEQTLRADEILPSESQKQDYBENSWDTSLSCTETVS	639
Db	334	PGKPSAFEPATEMQSVPNKALELKNEQTLRADEILPSESQKQDYBENSWDTSLSCTETVS	393
Qy	640	QKDVLCPKAAHQEKIDKINGKLEGPVPKDGILLKANCQMKVSIPTKALELMDMOTFKAAP	699
Db	394	QKDVLCPKAAHQEKIDKINGKLEGPVPKDGILLKANCQMKVSIPTKALELMDMOTFKAAP	453
Qy	700	EXPSAFEPATEMQSVPNKALELKNEQTLRADEILPSESQKQDYBENSWDSESICETVSQ	759
Db	454	EXPSAFEPATEMQSVPNKALELKNEQTLRADEILPSESQKQDYBENSWDSESICETVSQ	513
Qy	760	KDVCLPKATHQEKIDKINGKLEESPNDGFIKAPCRMKVSIPTKALELMDMOTFKAAPPE	819
Db	514	KDVCLPKATHQEKIDKINGKLEESPNDGFIKAPCRMKVSIPTKALELMDMOTFKAAPPE	573
Qy	820	KPSAFEPATEMQSVPNKALELKNEQTLRADQMPFPSSKQKDYBENSWDSESIRETVSQK	879
Db	574	KPSAFEPATEMQSVPNKALELKNEQTLRADQMPFPSSKQKDYBENSWDSESIRETVSQK	633
Qy	880	DVCPKATHQEKMDKISGKLEDSTSLKILDTVHSCERARELOKDHCEQRTGKWEQMKK	939
Db	634	DVCPKATHQEKMDKISGKLEDSTSLKILDTVHSCERARELOKDHCEQRTGKWEQMKK	693
Qy	940	FCVLKKLSEAKEIKSQLENOKVQKWEQELCSVR-LTL	975
Db	694	FCVLKKLSEAKEIKSQLENOKVQKWEQELCSVRFLTL	730

```

RESULT 15
US-09-834-759-494
; Sequence 494, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variant
; LOCATION: (1)...(743)
; OTHER INFORMATION: Xaa = Any amino acid

```

Query Match	47.0%; Score 3252.5; DB 4; Length 743;
Best Local Similarity	98.9%; Pred. No. 8.5e-215;
Matches 630; Conservative	2; Mismatches 4; Indels 1; Gaps 1;
Qy 340	MSPAKTSSEKPTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTVLEKGRSMAIACPTK 399
Db	
Qy 94	MSPAKETSEKPTWAAKGRPRKIAWEKKEPTVKTGCVARVTSNKTVLEKGRSMAIACPTK 153
Db	
Qy 400	ESSTKASANDORPSPESKOEEDSEYSCDSRSILPSSAKIQVCIPESIIYOKWWEINREVEE 459
Db	
Qy 154	ESSTKASANDORPSPESKOEEDSEYSCDSRSILPSSAKIQVCIPESIIYOKWWEINREVEE 213
Db	
Qy 460	PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPNFPPESSKQKDYEESSWDSSESICETV 519
Db	
Qy 214	PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPNFPPESSKQKDYEESSWDSSESICETV 273
Db	
Qy 520	SQKDVLCPKATHQKEIDKINGKLEESPNDGLLKATCGMKVSIPTKALELMDMQTFKAEP 579
Db	
Qy 274	SQKDVLCPKATHQKEIDKINGKLEESPNDGLLKATCGMKVSIPTKALELMDMQTFKAEP 333
Db	
Qy 580	PGPSPAFEPATEMQSVPNKALELKNEQTRADEIILPSESKQKDYEESSWDSSESICETVS 639
Db	
Qy 334	PGPSPAFEPATEMQSVPNKALELKNEQTLRADEIILPSESKQKDYEESSWDSSESICETVS 393
Db	
Qy 640	QKDVCIPKAAHQKEIDKINGKLESGPVKQGLLKANGCMKVISIPTKALELMDMQTFKAEP 699
Db	
Qy 394	QKDVCIPKAAHQKEIDKINGKLESGPVKQGLLKANGCMKVISIPTKALELMDMQTFKAEP 453
Db	
Qy 700	EKPSAFEPATEMQSVPNKALELKNEQTLRADBIILPSESKQKDYEESSWDSSESICETVSQ 759
Db	
Qy 454	EKPSAFEPATEMQSVPNKALELKNEQTLRADBIILPSESKQKDYEESSWDSSESICETVSQ 513
Db	
Qy 760	KDYCLPKATHQKEIDKINGKLEESPNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPBE 819
Db	
Qy 514	KDYCLPKATHQKEIDKINGKLEESPNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPBE 573
Db	
Qy 820	KPSAFEPATEMQSVPNKALELKNEQTLRADQMPFSESKQKDYEESSWDSSESIRETVSOK 879
Db	
Qy 574	KPSAFEPATEMQSVPNKALELKNEQTLRADQMPFSESKQKDYEESSWDSSESIRETVSOK 633
Db	
Qy 880	DVCVPKATHQKEMDKISGKLEDSTSLSKILDTVHSCERARELQKHCEORTGWEQMKKK 939
Db	
Qy 634	DVCVPKATHQKEMDKISGKLEDSTSLSKILDTVHSCERARELQKHCEORTGWEQMKKK 693
Db	
Qy 940	FCVLKKKLSGSAEIKSQLENQKQKWEQELCSVR-LTL 975
Db	
Qy 694	FCVLKKKLSGSAEIKSQLENQKQKWEQELCSVRFLT 730
Db	

Search completed: February 19, 2005, 02:13:30
Job time : 40.5216 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2005, 01:52:33 ; Search time 146.087 Seconds
(without alignments)
3550.262 Million cell updates/sec

Title: US-09-602-362E-23

Perfect score: 6320

Sequence: 1 MTRKKKTINLIQAQKRTA.....NHLKNRIYQVEKEKAETENS 1341

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Genesep1980s:*
- 2: Genesep1990s:*
- 3: Genesep2000s:*
- 4: Genesep2001s:*
- 5: Genesep2002s:*
- 6: Genesep2003as:*
- 7: Genesep2003bs:*
- 8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6920	100.0	1341	4	AAB84702 Amino aci
2	6920	100.0	1341	5	Abj05537 Breast ca
3	6920	100.0	1341	6	Abj37784 Human tum
4	6920	100.0	1341	6	Abj47548 Breast ca
5	6920	100.0	1341	7	Adl93227 Human bre
6	6892	99.6	1349	6	Abj37788 Human tum
7	6892	99.6	1349	7	Adl93235 Human bre
8	5130	74.1	1002	4	AAB84702 Human bre
9	5130	74.1	1002	5	Abg78918 Human bre
10	5130	74.1	1002	6	Abj37741 Human tum
11	5130	74.1	1002	7	Adl93137 Human bre
12	5130	74.1	1002	8	Adl93137 Human bre
13	5130	74.1	1095	4	AAB84702 Human bre
14	5130	74.1	1095	5	Abg78924 Human bre
15	5130	74.1	1095	6	Abj37747 Human tum
16	5130	74.1	1095	7	Adl93155 Human bre
17	5124	74.0	1013	6	Abj37783 Human tum
18	5124	74.0	1013	7	Adl93215 Human bre
19	3945	57.0	1239	6	Abj37789 Human tum
20	3945	57.0	1239	7	Adl93239 Human bre
21	3721	53.8	1225	6	Abj47547 Breast ca
22	3256.5	47.1	661	6	Abj37782 Human tum
23	3256.5	47.1	661	7	Adl93214 Human bre
24	3252.5	47.0	650	4	AAB50263 Human bre
25	3252.5	47.0	650	4	AAG65983 B726P spl

ALIGNMENTS

RESULT 1

AAB84702
ID AAB84702 standard; protein; 1341 AA.

XX AC AAB84702;

XX DT 17-SEP-2001 (first entry)

XX DE Amino acid sequence of a human cancer associated antigen.

XX KW Cancer associated antigen; INGI; tumour suppressor; cancer; vaccine.

XX OS Homo sapiens.

XX PN WO200147959-A2.

XX PD 05-JUL-2001.

XX PF 29-NOV-2000; 2000WO-US042334.

XX PR 30-NOV-1999; 99US-00451739.

XX PR 24-OCT-2000; 2000US-00602362.

XX PA (LUDW-) LUDWIG INST CANCER RES.
(SLOK) SLOAN KETTERING INST CANCER RES.
(CORR) CORNELL RES FOUND INC.

XX FI Jager D, Stockert E, Scanlan M, Knuth A, Old L, Gure A, Chen Y;

XX DR WPI; 2001-441706/47.

XX PT Isolated cancer associated nucleic acid molecule identified by SEREX
(serological identification of antigens by recombinant expression
cloning) technique, useful in nucleic acid based therapies to treat
cancer.

XX PS Claim 83; Page 53-57; 62pp; English.

XX CC The present sequence represents a human cancer associated antigen. The
sequence was identified using probes derived from the INGI gene. The INGI
gene is a tumour suppressor candidate gene. The cancer associated antigen
polynucleotides and polypeptides are useful for screening for the
possible presence of a pathological condition in a subject such as
cancer. The cancer associated antigen polypeptides are useful for
producing vaccines

XX SQ Sequence 1341 AA;

Aau33346 Human bre
Abg78913 Human bre
Abj37736 Human tum
Adl93131 Human bre
Ade44421 Human bre
Aau33358 Human bre
Abg78925 Human bre
Abj37748 Human tum
Adl93156 Human bre
Aab84703 Amino aci
Adj69751 Human hea
Aab84701 Amino aci
Aab50248 Human bre
Aag65986 B726P spl
Aau33349 Human bre
Abj37816 Human bre
Abj37739 Human tum
Adl93134 Human bre
Ade44424 Human bre
Aab50249 Human bre

26 3252.5 47.0 650 4 AAU33346
27 3252.5 47.0 650 5 ABG78913
28 3252.5 47.0 650 6 ABJ37736
29 3252.5 47.0 650 7 ADL93131
30 3252.5 47.0 650 8 ADE44421
31 3252.5 47.0 743 4 AAU33358
32 3252.5 47.0 743 5 ABG78925
33 3252.5 47.0 743 6 ABJ37748
34 3252.5 47.0 743 7 ADL93156
35 3075 44.4 1011 4 AAB84703
36 3075 44.4 1011 7 ADJ69751
37 2589 37.4 512 4 AAB84701
38 2270 32.8 466 4 AAB50248
39 2270 32.8 466 4 AAG65986
40 2270 32.8 466 4 AAU33349
41 2270 32.8 466 5 ABG78916
42 2270 32.8 466 6 ABJ37739
43 2270 32.8 466 7 ADL93134
44 2270 32.8 466 8 ADE44424
45 2256 32.6 445 4 AAB50249

Query Match 100.0%; Score 6920; DB 4; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTKRKTINLNIQDAQKTAHLWACVNGHEEVVTFVLDVKQDQDVLVDGEHRTPLMKALQC 60
DB 1 MTKRKTINLNIQDAQKTAHLWACVNGHEEVVTFVLDVKQDQDVLVDGEHRTPLMKALQC 60

QY 61 HOECANLIDSGADINLDVYGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPL 120
DB 61 HOECANLIDSGADINLDVYGNMALHYAVYSEILSVAKLISHGAVIEVHNKASLTPL 120

QY 121 LLSITKRSQIEVFLLIKNANANAVNKYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180
DB 121 LLSITKRSQIEVFLLIKNANANAVNKYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180

QY 181 CGVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPEGSTAGTDEAAPLAERTPTDAE 240
DB 181 CGVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPEGSTAGTDEAAPLAERTPTDAE 240

QY 241 SLVEKTPDEAAPLAVERTPTDAESLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSABE 300
DB 241 SLVEKTPDEAAPLAVERTPTDAESLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSABE 300

QY 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRK 360
DB 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRK 360

QY 361 IAWKKEPTPVKTGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORPPSPESKQEE 420
DB 361 IAWKKEPTPVKTGCVARVTSNKTIVLEKGRSKMIACPTKESSTKASANDORPPSPESKQEE 420

QY 421 DEEYSCDSRSLEPSSAKTOVCIPESYOKVMEINREVEEPKPSAFPAIEMQNSVFNK 480
DB 421 DEEYSCDSRSLEPSSAKTOVCIPESYOKVMEINREVEEPKPSAFPAIEMQNSVFNK 480

QY 481 APELKNEQTLRADPMFPFSPESKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKING 540
DB 481 APELKNEQTLRADPMFPFSPESKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKING 540

QY 541 KLEESPNDGGLLKATCGMKVSIPTKALEIKDMQTFKAEPPKPSAFPAIEMQNSVFNKA 600
DB 541 KLEESPNDGGLLKATCGMKVSIPTKALEIKDMQTFKAEPPKPSAFPAIEMQNSVFNKA 600

QY 601 LELKNEQTLRADEILPSPESKQDYENSWSDESICETVSQKDVCLPKAAHOKEDIDKING 660
DB 601 LELKNEQTLRADEILPSPESKQDYENSWSDESICETVSQKDVCLPKAAHOKEDIDKING 660

QY 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPKPSAFPAIEMQNSVFNKAL 720
DB 661 LEGSPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPKPSAFPAIEMQNSVFNKAL 720

QY 721 ELKNEQTLRADEILPSPESKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGKL 780
DB 721 ELKNEQTLRADEILPSPESKQDYENSWSDESICETVSQKDVCLPKATHQKEIDKINGKL 780

QY 781 EESPNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPPKPSAFPAIEMQNSVFNKALE 840
DB 781 EESPNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPPKPSAFPAIEMQNSVFNKALE 840

QY 841 LKNEQTLRADQMFPPSPESKQDYENSWSDESIRETVSQKDVCPKATHQKEMDKISGKLE 900
DB 841 LKNEQTLRADQMFPPSPESKQDYENSWSDESIRETVSQKDVCPKATHQKEMDKISGKLE 900

QY 901 DSTLSLSKILDTVHSCERARELOKHCEORTGMEQMKKFCVCLKKSEAKEIKSQLENQ 960
DB 901 DSTLSLSKILDTVHSCERARELOKHCEORTGMEQMKKFCVCLKKSEAKEIKSQLENQ 960

QY 961 KVKWEQELCSVRLTLNQBEERKNADILNEKIREELGRIEEHRKELEVKQOLEQALRIQ 1020
DB 961 KVKWEQELCSVRLTLNQBEERKNADILNEKIREELGRIEEHRKELEVKQOLEQALRIQ 1020

QY 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQKFNKYFEDI 1080

DB 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQKFNKYFEDI 1080
QY 1081 KILKEKNAELQMTLKLKEESLTKEASQYSGOLKVLIAENTMTLSKLKEKQDKEILEASIE 1140
DB 1081 KILKEKNAELQMTLKLKEESLTKEASQYSGOLKVLIAENTMTLSKLKEKQDKEILEASIE 1140
QY 1141 SHHPRLASAVODHQIIVTSRKSQBPAPHIAAGDACLQKMNVDVSTIYNNVHLHQPSEA 1200
DB 1141 SHHPRLASAVODHQIIVTSRKSQBPAPHIAAGDACLQKMNVDVSTIYNNVHLHQPSEA 1200
QY 1201 QRSKSLKININLYAGDALRENTLVSEHAQRDQRTQCMKEAHHYQNEQDNVKNHTSQ 1260
DB 1201 QRSKSLKININLYAGDALRENTLVSEHAQRDQRTQCMKEAHHYQNEQDNVKNHTSQ 1260
QY 1261 ESLQKFLQLOSQKNNMLQOQLVHAHKADNKSKITIDHFLERKMOHLLKKEKNEEIPNY 1320
DB 1261 ESLQKFLQLOSQKNNMLQOQLVHAHKADNKSKITIDHFLERKMOHLLKKEKNEEIPNY 1320
QY 1321 NNHLKNRIYQYEKEKAETENS 1341
DB 1321 NNHLKNRIYQYEKEKAETENS 1341

RESULT 2
ABJ05537
ID ABJ05537 standard; protein; 1341 AA.
XX
AC ABJ05537;
XX
DT 14-NOV-2002 (first entry)
XX
DE Breast cancer-associated protein 2.
XX
KW Breast cancer; breast cancer-associated gene sequence; drug development;
XX pharmacogenetics; biosensor development.
XX Unidentified.
XX WO200259377-A2.
XX
PD 01-AUG-2002.
XX
PF 24-JAN-2002; 2002WO-US002242;
XX
PR 24-JAN-2001; 2001US-0263965P.
PR 02-FEB-2001; 2001US-0265928P.
PR 09-APR-2001; 2001US-00829472.
PR 09-APR-2001; 2001US-0282698P.
PR 04-MAY-2001; 2001US-0288590P.
PR 29-MAY-2001; 2001US-0294443P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Gish KC, Afar D;
XX
FI WPI; 2002-583738/62.
XX
DR N-PSDB; ABT07694.
XX
PT Detecting a breast cancer-associated transcript in a patient's cell,
PT useful for diagnosing breast cancer, comprises contacting a biological
PT sample with a polynucleotide that selectively hybridizes with breast
PT cancer nucleic acids.
XX
PS Disclosure; Page 348-349; 414pp; English.
XX
CC The invention comprises a method of detecting a breast cancer-associated
CC transcript in a cell from a patient. The method of the invention involves
CC contacting a biological sample from the patient with a nucleotide that
CC hybridizes to one of the 69 breast cancer-associated gene sequences shown
CC in the specification. The method of the invention is useful in the
CC diagnosis or prognosis of breast cancer, and for detecting genes that are
CC up or down-regulated in breast cancer cells. Genes identified by the

method of the invention can be used in diagnostic purposes and also as targets for screening or therapeutic compounds that modulate breast cancer (e.g. hormones or antibodies). Identification of genes that are over or under expressed in breast cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure and biosensor development. Amino acid sequences ABJ0536 - ABJ05604 represent the proteins encoded by the 69 breast cancer-associated genes of the invention

Query Match 100.0%; Score 6920; DB 5; Length 1341;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTRKKTININIQDAQKRTALHWACVNGHREVTFLVDRKCOLDVLDGEHRTPLMKALQC 60
Db 1 MTRKKTININIQDAQKRTALHWACVNGHREVTFLVDRKCOLDVLDGEHRTPLMKALQC 60

Qy 61 HQEACANILDSGADINLDVYGNMALHYAVYSEILSVAKLLSHGAVIEVHNKASLTPL 120
Db 61 HQEACANILDSGADINLDVYGNMALHYAVYSEILSVAKLLSHGAVIEVHNKASLTPL 120

Qy 121 LLSITKRSQEIPEFLLIKNNANANVKKCTALMLAVCHGSSEIVGMLLQONVDFPAADI 180
Db 121 LLSITKRSQEIPEFLLIKNNANANVKKCTALMLAVCHGSSEIVGMLLQONVDFPAADI 180

Qy 181 CGVTAHYAVTCGPHHHEQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPTDAE 240
Db 181 CGVTAHYAVTCGPHHHEQIMEYIRKLSKNHONTNPEGTSAGTPDEAAPLAERTPTDAE 240

Qy 241 SLVEKTPDEAAPLVERPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQSAEE 300
Db 241 SLVEKTPDEAAPLVERPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKFEQSAEE 300

Qy 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360
Db 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWPAKGRPRK 360

Qy 361 IAMEKKEPTVKTCVARVTSNKTIVLEKGRSKMTACPTKESSTKASANDORFPSESQOE 420
Db 361 IAMEKKEPTVKTCVARVTSNKTIVLEKGRSKMTACPTKESSTKASANDORFPSESQOE 420

Qy 421 DEEYSCDSRSLFESSAKIQVCPESIQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480
Db 421 DEEYSCDSRSLFESSAKIQVCPESIQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480

Qy 481 APELKNQTLRADPMFPPEKQKDYENSWSLSECTVSKQVCLPKATHQKEIDKING 540
Db 481 APELKNQTLRADPMFPPEKQKDYENSWSLSECTVSKQVCLPKATHQKEIDKING 540

Qy 541 KLESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVPNKA 600
Db 541 KLESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVPNKA 600

Qy 601 LELKNEOTWRADEILPESKQKDYENSWSLSECTVSKQVCLPKAAHOKETDKINGK 660
Db 601 LELKNEOTWRADEILPESKQKDYENSWSLSECTVSKQVCLPKAAHOKETDKINGK 660

Qy 661 LEGSPVKDGLLKANCGRMKVSIPTKALELMDMTFKAEPPPEKPSAFEPATEMOKSVPNKAL 720
Db 661 LEGSPVKDGLLKANCGRMKVSIPTKALELMDMTFKAEPPPEKPSAFEPATEMOKSVPNKAL 720

Qy 721 ELKNEOTLRADDEILPESKQKDYENSWSLSECTVSKQVCLPKATHQKEIDKINGKL 780
Db 721 ELKNEOTLRADDEILPESKQKDYENSWSLSECTVSKQVCLPKATHQKEIDKINGKL 780

Qy 781 EESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPPEKPSAFEPATEMOKSVPNKALE 840
Db 781 EESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPPEKPSAFEPATEMOKSVPNKALE 840

Qy 841 LKNEQTLRADQMFPPESEKQKVEENSWDSLSLRETQVSKQVCPKATHQKEMDKISGKLE 900

Db 841 LKNEQTLRADQMFPPESEKQKVEENSWDSLSLRETQVSKQVCPKATHQKEMDKISGKLE 900

Qy 901 DSTSLSKILDTVHSCERARELQKDHCHQRTQKMEQMKKFCVLKKLSEAKEIKSOLBNQ 960

Db 901 DSTSLSKILDTVHSCERARELQKDHCHQRTQKMEQMKKFCVLKKLSEAKEIKSOLBNQ 960

Qy 961 KVKWEQELCSVRLTLNQEEERKRNADILNEKIREBELGRIBEQHKELEVKQOLEQALRIQ 1020

Db 961 KVKWEQELCSVRLTLNQEEERKRNADILNEKIREBELGRIBEQHKELEVKQOLEQALRIQ 1020

Qy 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKIAMLKLEIATLKHQYQEKENKYFEDI 1080

Db 1021 DIELKSVESNLNQVSHTHENENYLLHENCMLKKIAMLKLEIATLKHQYQEKENKYFEDI 1080

Qy 1081 KILKEKNAELQMTLKLKEESLTKRASQVSGOLKVLIAENTMLTSLKLEKQDKLELEAEIE 1140

Db 1081 KILKEKNAELQMTLKLKEESLTKRASQVSGOLKVLIAENTMLTSLKLEKQDKLELEAEIE 1140

Qy 1141 SHHPRLASAVQDHDQIIVTSRKSQBPAPHIAGDACLQRMNVDSSTIYNNEVLHQPLSEA 1200

Db 1141 SHHPRLASAVQDHDQIIVTSRKSQBPAPHIAGDACLQRMNVDSSTIYNNEVLHQPLSEA 1200

Qy 1201 QRKSKSLKINLYAGDALRENTLVSEHAQDQRETQCMKEAHEMYQNEQDNVKNHTSQO 1260

Db 1201 QRKSKSLKINLYAGDALRENTLVSEHAQDQRETQCMKEAHEMYQNEQDNVKNHTSQO 1260

Qy 1261 ESLDQKLPQLQSKNMVLOQQLVHAHKADNKSKITIDHFLERKNQHHLLKEKNEEIEFNY 1320

Db 1261 ESLDQKLPQLQSKNMVLOQQLVHAHKADNKSKITIDHFLERKNQHHLLKEKNEEIEFNY 1320

Qy 1321 NNHLKNRIYQYEKEKAETENS 1341

Db 1321 NNHLKNRIYQYEKEKAETENS 1341

RESULT 3
ABJ37784
ID ABJ37784 standard; protein; 1341 AA.

XX AC ABJ37784;
XX AC
XX DT 15-MAY-2003 (first entry)
XX DE Human tumour-related protein - SEQ ID No 565.
XX KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
XX KW tumour; breast cancer; cancer; immune response stimulation.
XX OS Homo sapiens.
XX PN WO200283956-A1.
XX PD 24-OCT-2002.
XX PF 15-APR-2002; 2002WO-US012378.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX PA (CORI-) CORIXA CORP.
XX PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
XX PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
XX PI Medvick TS, McNeill PD, Durham M;
XX DR WPI; 2003-103376/09.
XX PT New polypeptide and polynucleotide useful for stimulating and/or
XX PT expanding T cells specific for a tumor protein and treating breast
XX PT cancer.

PS	Example 9; Page 342-346; 375pp; English.	
XX	The invention comprises a method of stimulating and/or expanding T cells specific for a tumour protein. The invention further comprises human nucleic acids and proteins that are associated with tumours (e.g. breast cancer). The method and sequences of the invention are useful for stimulating and/or expanding T cells specific for a tumour protein, detecting the presence of cancer, stimulating an immune response in a patient and treating breast cancer. The present amino acid sequence represents a human tumour-related protein	
CC	Sequence 1341 AA;	
CC	Query Match 100.0%; Score 6920; DB 6; Length 1341;	
CC	Best Local Similarity 100.0%; Pred. No. 0;	
CC	Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MTKKKTINLNIQDAKRTALHWACVNGHEEVTFLVDRKQQLDVLGDEHRTPLMKALQC 60	
DB	1 MTKKKTINLNIQDAKRTALHWACVNGHEEVTFLVDRKQQLDVLGDEHRTPLMKALQC 60	
QY	61 HOEACANILDSGADINLVYGNWALHYAVYSILSVVAKLHSGAVIEVHNKASLTPL 120	
DB	61 HOEACANILDSGADINLVYGNWALHYAVYSILSVVAKLHSGAVIEVHNKASLTPL 120	
QY	121 LLSITKRSQIIVFELLKNANANVYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180	
DB	121 LLSITKRSQIIVFELLKNANANVYKCTALMLAVCHGSSEIVGMLLQONVDVFAADI 180	
QY	181 CGVTAHVAVTCGPHHIEQIMYIRKLSKNHNTNPEGSTAGTPDEAALPRTPTAE 240	
DB	181 CGVTAHVAVTCGPHHIEQIMYIRKLSKNHNTNPEGSTAGTPDEAALPRTPTAE 240	
QY	241 SLVEKTPDEAAPLVERTPDAESLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSABE 300	
DB	241 SLVEKTPDEAAPLVERTPDAESLVEKTPDEAASLVEGTSKIOCLEKATSGKFEQSABE 300	
QY	301 TPREITSPAKETSEKFTWPAKGRPKIAWEKEDTPREIMSPAKETSEKFTWAAKGRPK 360	
DB	301 TPREITSPAKETSEKFTWPAKGRPKIAWEKEDTPREIMSPAKETSEKFTWAAKGRPK 360	
QY	361 IAWKKEKTPVKTGCVARTSNKTKVLEKGRSKMIACPTKESSTKASANDORPPSPKOE 420	
DB	361 IAWKKEKTPVKTGCVARTSNKTKVLEKGRSKMIACPTKESSTKASANDORPPSPKOE 420	
QY	421 DEEYSCDSRSLPSSAKIQVCIPESIQYKWEINREVEEPPKPSAPKPAIEMQNSVPNK 480	
DB	421 DEEYSCDSRSLPSSAKIQVCIPESIQYKWEINREVEEPPKPSAPKPAIEMQNSVPNK 480	
QY	481 APELKNQOTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKING 540	
DB	481 APELKNQOTLRADPMPFPESKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKING 540	
QY	541 KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVPNKA 600	
DB	541 KLEESPNKDGLLKATCGMKVSIPTKALELMDMTFKAEPPGKPSAFEPATEMOKSVPNKA 600	
QY	601 LELKNEQTRADPEILPSSKQDYENSWSDESICETVSKQDVCLPKAAHQKEIDKINGK 660	
DB	601 LELKNEQTRADPEILPSSKQDYENSWSDESICETVSKQDVCLPKAAHQKEIDKINGK 660	
QY	661 LEGSPVKDGLLKANCCKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKAL 720	
DB	661 LEGSPVKDGLLKANCCKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKAL 720	
QY	721 ELKNEQTLRADPEILPSSKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKINGKL 780	
DB	721 ELKNEQTLRADPEILPSSKQDYENSWSDESICETVSKQDVCLPKATHQKEIDKINGKL 780	
QY	781 EESPNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKALE 840	
DB	781 EESPNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPATEMOKSVPNKALE 840	

QY	841 LKNEQTLRADOMPSPSSKQKVVENSWSDESISLRTVSKQDVCPKATHQKEMDKISGLE 900	
DB	841 LKNEQTLRADOMPSPSSKQKVVENSWSDESISLRTVSKQDVCPKATHQKEMDKISGLE 900	
QY	901 DSTSLSKILDVTHSCBRARELQKHCEQRTQKWEQMKKFCVLKKLSEAKEIKSQLENQ 960	
DB	901 DSTSLSKILDVTHSCBRARELQKHCEQRTQKWEQMKKFCVLKKLSEAKEIKSQLENQ 960	
QY	961 KVKWEQELCSVRLTNOEEREKRNADILNEKIREELGRIEEQHRKELEVQKQLEQALRIQ 1020	
DB	961 KVKWEQELCSVRLTNOEEREKRNADILNEKIREELGRIEEQHRKELEVQKQLEQALRIQ 1020	
QY	1021 DIELKSVESNLQVSHTHENENYLLHENCMKKEIAMLKLEIATLKHQYQEKENKYPEDI 1080	
DB	1021 DIELKSVESNLQVSHTHENENYLLHENCMKKEIAMLKLEIATLKHQYQEKENKYPEDI 1080	
QY	1081 KILKEKNAELQMTLKLKEESLTKRASQYSQQLVLAENTMTLSKLEKQDKETLEAEIE 1140	
DB	1081 KILKEKNAELQMTLKLKEESLTKRASQYSQQLVLAENTMTLSKLEKQDKETLEAEIE 1140	
QY	1141 SHHPRLASAVODHDOIIVTSRKSOEPAPHIAGDACLQRMNVDSSTIYNNVHLQPLSEA 1200	
DB	1141 SHHPRLASAVODHDOIIVTSRKSOEPAPHIAGDACLQRMNVDSSTIYNNVHLQPLSEA 1200	
QY	1201 QRKSKSLKINLVAGDALRENTLVSEHAQDRETOCQMKAEAHMYQNEQDNVNVKHTEQQ 1260	
DB	1201 QRKSKSLKINLVAGDALRENTLVSEHAQDRETOCQMKAEAHMYQNEQDNVNVKHTEQQ 1260	
QY	1261 ESLDQKLFQLOSKNMWLOQOLVHAHKKADNKSITTDIHFELERKMOHLLKKNKEEIFNY 1320	
DB	1261 ESLDQKLFQLOSKNMWLOQOLVHAHKKADNKSITTDIHFELERKMOHLLKKNKEEIFNY 1320	
QY	1321 NNHLKNRIYQYEKEKASTENS 1341	
DB	1321 NNHLKNRIYQYEKEKASTENS 1341	
XX	RESULT 4	
XX	ABR47548	
ID	ABR47548 standard; protein; 1341 AA.	
XX	AC ABR47548;	
XX	DT 12-JUN-2003 (first entry)	
DE	Breast cancer associated protein sequence SEQ ID NO:334.	
XX	Human; breast cancer; cytostatic; gene therapy.	
OS	Homo sapiens.	
XX	WO2003004989-A2.	
XX	16-JAN-2003.	
XX	21-JUN-2002; 2002WO-US019669.	
XX	21-JUN-2001; 2001US-0299887P.	
XX	27-JUN-2001; 2001US-0301572P.	
XX	18-JUL-2001; 2001US-0306501P.	
XX	25-SEP-2001; 2001US-0325002P.	
XX	05-MAR-2002; 2002US-0362585P.	
XX	14-MAY-2002; 2002US-0380391P.	
XX	(MILL-) MILLENIUM PHARM INC.	
XX	Lillie J, Gannavarapu M, Glatt K, Hoersh S, Kamatkar S;	
PI	Mertens M, Monahan JE, Myer Y, Wang Y, Xu Y, Zhao X, Meyers RE;	
XX	Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;	
DR	WPI: 2003-210381/20.	
DR	N-PSDB; ACC50246.	
XX		

PT Breast cancer diagnosis or treatment by comparing the level of expression
 PT of a marker in a patient sample with that in the control non-breast
 PT cancer sample.

XX Claim 1; SEQ ID NO 334; 128pp; English.

XX The present invention describes a method for assessing whether a patient
 CC is afflicted with breast cancer. The method comprises comparing the level
 CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
 CC ABR47386 to ABR47632) in a patient sample and the normal level of
 CC expression of the marker in a control non-breast cancer sample, where a
 CC significant increase in the level of expression of the marker in the
 CC patient sample and the normal level is an indication that the patient is
 CC afflicted with breast cancer. The breast cancer associated sequences from
 CC the present invention have cytostatic activities and can be used in gene
 CC therapy. The method is useful for diagnosing and treating breast cancer.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1341 AA;

Query Match 100.0%; Score 6920; DB 6; Length 1341;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTKRKTINLNIOAQRKTALHWACVNGHEEVVTLVDRKQDLVDLGEHRTPLMKALQC 60
 Db 1 MTKRKTINLNIOAQRKTALHWACVNGHEEVVTLVDRKQDLVDLGEHRTPLMKALQC 60

Qy 61 HQEACANILDSGADINLVGYGNALHYAVYSILSVVAKLLSHGAVIEVHNKASLTPL 120
 Db 61 HQEACANILDSGADINLVGYGNALHYAVYSILSVVAKLLSHGAVIEVHNKASLTPL 120

Qy 121 LLSITKSEQIVPELLIKNANANAVKYKCTALMLAVCHGSSEIVGMLLQNVDFPAADI 180
 Db 121 LLSITKSEQIVPELLIKNANANAVKYKCTALMLAVCHGSSEIVGMLLQNVDFPAADI 180

Qy 181 CGVTAHYAVTCGPHHIEQIMEYIRKLSKHQNTNPEGTSAGTPDEAALPAERTPTDAE 240
 Db 181 CGVTAHYAVTCGPHHIEQIMEYIRKLSKHQNTNPEGTSAGTPDEAALPAERTPTDAE 240

Qy 241 SLVEKTPDEAALVETPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKPEQSAEE 300
 Db 241 SLVEKTPDEAALVETPTDAESLVEKTPDEAASLVEGTSKIQCLEKATSGKPEQSAEE 300

Qy 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWAAKGRPRK 360
 Db 301 TPREITSPAKETSEKFTWPAKGRPRKIAWEKEDTPREIMSPAKETSEKFTWAAKGRPRK 360

Qy 361 IAWKKEKTPVKTGCVARVTSNKTVLEKGRSKMIACPTKESSTKASANDORPPSESQOE 420
 Db 361 IAWKKEKTPVKTGCVARVTSNKTVLEKGRSKMIACPTKESSTKASANDORPPSESQOE 420

Qy 421 DEEYSCDSRSLFESSAKIQVIPSIIYQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480
 Db 421 DEEYSCDSRSLFESSAKIQVIPSIIYQKWEINREVEEPPKPSAFKPAIEMQNSVPNK 480

Qy 481 AFEKNEQTLRADMPFPESKQDYENSWDSLSLCEVTSQKDVCLPKATHQKEIDKING 540
 Db 481 AFEKNEQTLRADMPFPESKQDYENSWDSLSLCEVTSQKDVCLPKATHQKEIDKING 540

Qy 541 KLESPNKGILLKATCGMKVSIPTKALELMDQTFKAEPPGKPSAFEPATEMOKSVPNKA 600
 Db 541 KLESPNKGILLKATCGMKVSIPTKALELMDQTFKAEPPGKPSAFEPATEMOKSVPNKA 600

Qy 601 LELKNEQTLRADEILPSESQKQDYENSWDTSLSLCEVTSQKDVCLPKAAHQKEIDKING 660
 Db 601 LELKNEQTLRADEILPSESQKQDYENSWDTSLSLCEVTSQKDVCLPKAAHQKEIDKING 660

Qy 661 LEGSPVKDGLLKANCGRMKVSIPTKALELMDQTFKAEPPKPSAFEPATEMOKSVPNKAL 720
 Db 661 LEGSPVKDGLLKANCGRMKVSIPTKALELMDQTFKAEPPKPSAFEPATEMOKSVPNKAL 720

Qy 721 ELKNEQTLRADEILPSESQKQDYENSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 780
 Db 721 ELKNEQTLRADEILPSESQKQDYENSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 780

Qy 781 BESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPKPSAFEPATEMOKSVPNKALE 840
 Db 781 BESPDNDGFLKAPCRMKVSIPTKALELMDQTFKAEPPKPSAFEPATEMOKSVPNKALE 840

Qy 841 LKNEQTLRADOMPFPESKQKQDYENSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 900
 Db 841 LKNEQTLRADOMPFPESKQKQDYENSWDSLSLCEVTSQKDVCLPKATHQKEIDKINGKL 900

Qy 901 DSTLSLSKILDTVHSCERARELQKDHCEQRTCKMEQMKKFCVLLKKLSEAKEIKSOLBNQ 960
 Db 901 DSTLSLSKILDTVHSCERARELQKDHCEQRTCKMEQMKKFCVLLKKLSEAKEIKSOLBNQ 960

Qy 961 KVKWEQELCSVRLTLNQEEKRRNADILNEKIRELGRIBEQHRKELBVKQOLEQALRIQ 1020
 Db 961 KVKWEQELCSVRLTLNQEEKRRNADILNEKIRELGRIBEQHRKELBVKQOLEQALRIQ 1020

Qy 1021 DIEKLSVESNLNOYSHTHENENYLLHENCMLKKEIAMLKLEIATILKHQYQEKENKVFEDI 1080
 Db 1021 DIEKLSVESNLNOYSHTHENENYLLHENCMLKKEIAMLKLEIATILKHQYQEKENKVFEDI 1080

Qy 1081 KILKEKNAELQMTLKLKEESLTKEASQVSGQLKVLIAENTMTLSKLEKQDKKEILEABIE 1140
 Db 1081 KILKEKNAELQMTLKLKEESLTKEASQVSGQLKVLIAENTMTLSKLEKQDKKEILEABIE 1140

Qy 1141 SHHPRLASAVODHDQIVTSRKQBPAPHIAAGDACLQKQNVDSSTIYNNNEVLHQPLSEA 1200
 Db 1141 SHHPRLASAVODHDQIVTSRKQBPAPHIAAGDACLQKQNVDSSTIYNNNEVLHQPLSEA 1200

Qy 1201 ORKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAHHYQNEQNVNKHQTSQQ 1260
 Db 1201 ORKSKSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAHHYQNEQNVNKHQTSQQ 1260

Qy 1261 ESLDQKFLQSQKNMVLQQLVHAHKKADNKSKITIDHFLERKMQHLLKKEKNEEIFYN 1320
 Db 1261 ESLDQKFLQSQKNMVLQQLVHAHKKADNKSKITIDHFLERKMQHLLKKEKNEEIFYN 1320

Qy 1321 NNHLKNRIYQVEKEKAETENS 1341
 Db 1321 NNHLKNRIYQVEKEKAETENS 1341

RESULT 5
 ADL93227
 ID ADL93227 standard; protein; 1341 AA.
 XX AC ADL93227;
 XX DT 20-MAY-2004 (first entry)
 XX DE Human breast cancer-associated polypeptide #49.
 XX KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
 XX OS Homo sapiens.
 XX PN US2003166022-A1.
 XX PD 04-SEP-2003.
 XX PF 15-APR-2002;-2002US-00124805.
 XX PR 28-DEC-1998; 98US-00222575.
 XX PR 02-APR-1999; 99US-00285480.
 XX PR 23-JUN-1999; 99US-00339338.
 XX PR 02-SEP-1999; 99US-00389681.
 XX PR 03-NOV-1999; 99US-00433826.
 XX PR 17-APR-2000; 2000US-00551621.
 XX PR 08-JUN-2000; 2000US-00590751.

PR	22-JUN-2000;	2000US-00604287.
PR	20-JUL-2000;	2000US-00620405.
PR	13-APR-2001;	2001US-00834759.
PR	07-DEC-2001;	2001US-00007805.
PR	13-FEB-2002;	2002US-00076622.
XX	(CORI-) CORIXA CORP.	
XX	Houghton RL, Sleath PR, Persing DH;	
XX	WPI; 2003-874918/81.	
DR	N-PSDB; ADL93226.	
XX	An isolated oncogenic polypeptide useful for preventing, diagnosing and	
PT	treating breast cancer.	
XX	Example 12; SEQ ID NO 565; 294pp; English.	
PS	The invention relates to an isolated breast cancer-associated	
CC	polypeptide. The polypeptide may be used for the diagnosis and treatment	
CC	of breast cancers. The methods are useful for detecting the presence of a	
CC	cancer in a patient and treating a cancer in a patient. The present	
CC	sequence represents the amino acid sequence of a human breast cancer-	
CC	associated polypeptide.	
XX	Sequence 1341 AA;	
SQ	Query Match 100.0%; Score 6920; DB 7; Length 1341;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MTKRKKTINLMIQDAQKRTALHWACVNGHEEVTFLVDRKCOLDVLGDEHRTPLMKALQC 60	
Db	1 MTKRKKTINLMIQDAQKRTALHWACVNGHEEVTFLVDRKCOLDVLGDEHRTPLMKALQC 60	
Qy	61 HOEACANLIDSGADINLVDVYGNNALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120	
Db	61 HOEACANLIDSGADINLVDVYGNNALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 120	
Qy	121 LLSITKRSEIQVEFLLIKNNANANAANKYKCTALMLAVCHGSGSEIVGMLLQONVDVFAADI 180	
Db	121 LLSITKRSEIQVEFLLIKNNANANAANKYKCTALMLAVCHGSGSEIVGMLLQONVDVFAADI 180	
Qy	181 CGVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPGTSGAGTDEAAPLAERTPTDAE 240	
Db	181 CGVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPGTSGAGTDEAAPLAERTPTDAE 240	
Qy	241 SLVEKTPDEAAPLVERTPTDAESLVEKTPDEAASLVEGTSDKIQCLEKATSGKFFQSAAE 300	
Db	241 SLVEKTPDEAAPLVERTPTDAESLVEKTPDEAASLVEGTSDKIQCLEKATSGKFFQSAAE 300	
Qy	301 TPREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRK 360	
Db	301 TPREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPREIMSPAKETSEKFTWAAKGRPRK 360	
Qy	361 IAWEKKEPTVKTCGVARTVSNKTKVLEKGRSKWIACTPKESSTKASANDQRPFSSEKQBE 420	
Db	361 IAWEKKEPTVKTCGVARTVSNKTKVLEKGRSKWIACTPKESSTKASANDQRPFSSEKQBE 420	
Qy	421 DEEYSCDSRLSPFESSAKIQVCIPESITYQKVMIEINREVEEPPKKPSAFKPAIEMQNSVPNK 480	
Db	421 DEEYSCDSRLSPFESSAKIQVCIPESITYQKVMIEINREVEEPPKKPSAFKPAIEMQNSVPNK 480	
Qy	481 AFELKNEQTLRADMPMPFPPSKQXDYEENSWDESLSCTETVSQKDVCILPKATHOKEIDKING 540	
Db	481 AFELKNEQTLRADMPMPFPPSKQXDYEENSWDESLSCTETVSQKDVCILPKATHOKEIDKING 540	
Qy	541 KLEESPNKQGLLKATCGMKVSIPTKALELKDQMTFKAEPPGKPSAFEPATEMQKSPVNKA 600	
Db	541 KLEESPNKQGLLKATCGMKVSIPTKALELKDQMTFKAEPPGKPSAFEPATEMQKSPVNKA 600	
Qy	601 LELKNEQTLWRADLILPSESKQXDYEENSWDTELSCTETVSQKDVCLPKAAHKEIDKING 660	

PR 02-APR-1999; 99US-00285480.
 PR 23-JUN-1999; 99US-00339338.
 PR 02-SEP-1999; 99US-00389681.
 PR 03-NOV-1999; 99US-00433826.
 PR 17-APR-2000; 2000US-00551821.
 PR 08-JUN-2000; 2000US-00590751.
 PR 22-JUN-2000; 2000US-00604287.
 PR 20-JUL-2000; 2000US-00620405.
 PR 13-APR-2001; 2001US-00834759.
 PR 07-DEC-2001; 2001US-00007805.
 PR 13-FEB-2002; 2002US-00076622.
 XX (CORI-) CORIXA CORP.
 XX
 XX Houghton RL, Sleath PR, Persing DH;
 XX
 DR WPI; 2003-874918/81.
 DR N-PSDB; ADL93228.
 XX
 PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
 PT treating breast cancer.
 XX
 PS Example 12; SEQ ID NO 573; 294pp; English.
 XX
 CC The invention relates to an isolated breast cancer-associated
 CC polypeptide. The polypeptide may be used for the diagnosis and treatment
 CC of breast cancers. The methods are useful for detecting the presence of a
 CC cancer in a patient and treating a cancer in a patient. The present
 CC sequence represents the amino acid sequence of a human breast cancer-
 CC associated polypeptide.
 XX
 SQ Sequence 1349 AA;

Query Match 99.6%; Score 6892; DB 7; Length 1349;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1336; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TKRKTINLNTQDAQKRTALHWCVNGHEEVVTVFLVDRKCOLDLVDGHRTPLMKALQCH 61
 DB 10 TKRKTINLNTQDAQKRTALHWCVNGHEEVVTVFLVDRKCPDVLVDGHRTPLMKALQCH 69
 QY 62 QEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 121
 DB 70 QEACANILIDSGADINLVDVYGNMALHYAVYSEILSVVAKLLSHGAVIEVHNKASLTPL 129
 QY 122 LSITKRSQIVFELLKNNANANVKNYKCTALMLAVCHGSSEIVGMLLQQNVDFVPAADIC 181
 DB 130 LSITKRSQIVFELLKNNANANVKNYKCTALMLAVCHGLSEIVGMLLQQNVDFVPAADIC 189
 QY 182 GVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPEGTSAGTDPDEAPLAERTPDTAES 241
 DB 190 GVTAEHYAVTCGFHHIHEQIMEYIRKLSKNHNTNPEGTSAGTDPDEAPLAERTPDTAES 249
 QY 242 LVEKTPDEAALPVERTPTASLSVEKTPDEAASLVEGTSKIQCLEKATSKGFSQABET 301
 DB 250 LVEKTPDEAALPVERTPTASLSVEKTPDEAASLVEGTSKIQCLEKATSKGFSQABET 309
 QY 302 PREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPTREIMSPAKETSEKFTWAAKGRPKI 361
 DB 310 PREITSPAKETSEKFTWPAKGRPKIAWEKKEDTPTREIMSPAKETSEKFTWAAKGRPKI 369
 QY 362 AWEKKETPVKTCGVARTVSNKTKVLEKGRSWMIACTPKTESSTKASANDQRPFSKQBED 421
 DB 370 AWEKKETPVKTCGVARTVSNKTKVLEKGRSWMIACTPKTESSTKASANDQRPFSKQBED 429
 QY 422 EYSCDSRSLFESSAKIQVCIPESYIQVKMEINREVEEPPKPSAPKPAIEMQNSVPNKA 481
 DB 430 EYSCDSRSLFESSAKIQVCIPESYIQVKMEINREVEEPPKPSAPKPAIEMQNSVPNKA 489
 QY 482 FELKNEOTLRADMPFPSPKOKDYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGK 541
 DB 490 FELKNEOTLRADMPFPSPKOKDYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGK 549

QY 542 LEESPNDGGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKAL 601
 DB 550 LEESPNDGGLLKATCGMKVSIPTKALELKMOTFKABPPGKPSAFEPATEMOKSVPNKAL 609
 QY 602 ELKNEQTRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKAAHQKEIDKINGKL 661
 DB 610 ELKNEQTRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKAAHQKEIDKINGKL 669
 QY 662 EGSPVKDGLLKANCCKMKVSIPTKALELMDMOTFKABPPGKPSAFEPATEMOKSVPNKALE 721
 DB 670 EGSPVKDGLLKANCCKMKVSIPTKALELMDMOTFKABPPGKPSAFEPATEMOKSVPNKALE 729
 QY 722 LKNEQTLRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGKLE 781
 DB 730 LKNEQTLRADEILPSESOKDYEENSWSDESLSCTVSKQDVCLPKATHOKEIDKINGKLE 789
 QY 782 ESPNDGFLKAPCKMKVSIPTKALELMDMOTFKABPPGKPSAFEPATEMOKSVPNKALEL 841
 DB 790 ESPNDGFLKAPCKMKVSIPTKALELMDMOTFKABPPGKPSAFEPATEMOKSVPNKALEL 849
 QY 842 KNEQTLRADQMPFPSESOKDYEENSWSDESLSCTVSKQDVCPKATHOKEIDKINGKLE 901
 DB 850 KNEQTLRADQMPFPSESOKDYEENSWSDESLSCTVSKQDVCPKATHOKEIDKINGKLE 909
 QY 902 STLSKLIDTVHSCERARELOKHCEORTGHMEOMKKFCVLKKLSEAKIKSOLENOK 961
 DB 910 STLSKLIDTVHSCERARELOKHCEORTGHMEOMKKFCVLKKLSEAKIKSOLENOK 969
 QY 962 VKWEOLCSVRLTUNQEEKRNADIINEKIREELGRIEEOHKELEVKQOLEALRIQD 1021
 DB 970 VKWEOLCSVRLTUNQEEKRNADIINEKIREELGRIEEOHKELEVKQOLEALRIQD 1029
 QY 1022 IELKSVESNLNQVSHTHENENYLLHENCMLKKEIATLKLHQTATLKHQYQEKENYFEDIK 1081
 DB 1030 IELKSVESNLNQVSHTHENENYLLHENCMLKKEIATLKLHQTATLKHQYQEKENYFEDIK 1089
 QY 1082 ILKEKNAELQMTLKLKESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIES 1141
 DB 1090 ILKEKNAELQMTLKLKESLTKRASQYSGQLKVLIAENTMLTSKLKEKODKEILEAEIES 1149
 QY 1142 HHPLASAVODHDOIVTSRSQOEPAFHITAGDACLQKMNVDVSTIYNNVHLHOPLSAQ 1201
 DB 1150 HHPLASAVODHDOIVTSRSQOEPAFHITAGDACLQKMNVDVSTIYNNVHLHOPLSAQ 1209
 QY 1202 RKSKSLINLNYAGDALRENTLVSEHAQDQRETQCMKEAHEMYQNEODNVNKHTEQOE 1261
 DB 1210 RKSKSLINLNYAGDALRENTLVSEHAQDQRETQCMKEAHEMYQNEODNVNKHTEQOE 1269
 QY 1262 SLQDKLFQLOSKNNWLOQLVHAHKADNKSKITIDHFLERKQHLLKEKNEIFNYN 1321
 DB 1270 SLQDKLFQLOSKNNWLOQLVHAHKADNKSKITIDHFLERKQHLLKEKNEIFNYN 1329
 QY 1322 NHLKNRIYQYEKEAETE 1339
 DB 1330 NHLKNRIYQYEKEAETE 1347
 RESULT 8
 AAU33351
 ID AAU33351 standard; protein; 1002 AA.
 XX
 AC AAU33351;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human breast cancer protein B726P from alternatively spliced cDNA.
 XX
 KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
 KW gene therapy; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200179286-A2.

XX PD 25-OCT-2001.
 XX PF 12-APR-2001; 2001WO-US012164.
 XX PR 17-APR-2000; 2000US-00551621.
 XX PR 08-JUN-2000; 2000US-00390751.
 XX PR 22-JUN-2000; 2000US-00604287.
 XX PR 20-JUL-2000; 2000US-00620405.
 XX PA (CORI-) CORIXA CORP.
 XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
 XX DR WPI; 2001-611721/70.
 XX DR N-PSDB; AAS47411.
 XX PT Breast Tumor Proteins and nucleic acids useful for the prevention,
 XX PT diagnosis and treatment of breast cancer.
 XX PS Claim 3; Page 281-283; 297pp; English.
 XX CC The invention relates to isolated breast tumour proteins and nucleic
 CC acids that encode them, including immunogenic fragments of the proteins.
 CC Also included are expression vectors expressing the proteins, transformed
 CC cells and antibodies raised against the proteins or an antigen presenting
 CC cell expressing the protein. The proteins and nucleic acids may be used
 CC in the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate breast tumour protein expression, i.e. breast tumours and
 CC breast cancer e.g. by gene therapy. The nucleic acids and their
 CC complements may also be used as DNA probes in diagnostic assays to detect
 CC and quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. The
 CC proteins, nucleic acids and antibodies may be used in assays to identify
 CC modulators (e.g. antagonists) of breast tumour protein expression and
 CC activity. The antibodies and antagonists may also be used to down
 CC regulate expression and activity. The antibodies may also be used as
 CC diagnostic agents for detecting the presence of the proteins in samples
 CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
 CC purification diagnostic techniques. The present sequence is a breast
 CC tumour protein encoded by a cDNA from a breast tumour cDNA library
 CC isolated by subtractive hybridisation against a normal breast cDNA
 XX library
 XX SQ Sequence 1002 AA;

Query Match 74.1%; Score 5130; DB 4; Length 1002;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAGRPRIAEKKEKTPVKTGCVARVTSNKTKEGSRKMIACPTK 399
 Db 1 MSPAKETSEKFTWAAGRPRIAEKKEKTPVKTGCVARVTSNKTKEGSRKMIACPTK 60

Qy 400 ESSTKASANDORPSESQEBDEYSCDSRLFPSSAKIQVICIPESIQKWEINREVEE 459
 Db 61 ESSTKASANDORPSESQEBDEYSCDSRLFPSSAKIQVICIPESIQKWEINREVEE 120

Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPPEPSKQDYBENSWSLCETV 519
 Db 121 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADMPPEPSKQDYBENSWSLCETV 180

Qy 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMTFKAEP 579
 Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMTFKAEP 240

Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETWADILPSESQKDYBENSWSLCETVS 639
 Db 241 PGKPSAFEPATEMOKSVPNKALELKNQETWADILPSESQKDYBENSWSLCETVS 300

Qy 640 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMTFKAEP 699
 Db 301 QKDVCLPKAAHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMTFKAEP 360

700 EKPSAFEPATEMOKSVPNKALELKNQETWADILPSESQKDYBENSWSLCETVSQ 759
 Db 361 EKPSAFEPATEMOKSVPNKALELKNQETWADILPSESQKDYBENSWSLCETVSQ 420

Qy 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMTFKAEP 819
 Db 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLLKATCGMKVSIPTKALELMDMTFKAEP 480

Qy 820 KPSAFEPATEMOKSVPNKALELKNQETWADILPSESQKDYBENSWSLCETVSQ 879
 Db 481 KPSAFEPATEMOKSVPNKALELKNQETWADILPSESQKDYBENSWSLCETVSQ 540

Qy 880 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMQMKKK 939
 Db 541 DVCVPKATHQKEMDKISGKLESDSTLSKILDTVHSCERARELQKDHCEQRTGKMQMKKK 600

Qy 940 FCVLKKKLESAKEIKSQLENOQVKEQELCSVRLTLNQEBEKKRRNADILNEKIRELGRI 999
 Db 601 FCVLKKKLESAKEIKSQLENOQVKEQELCSVRLTLNQEBEKKRRNADILNEKIRELGRI 660

Qy 1000 EEOHKELEVKQOLEQALRIQDIELKSVEINLVSHNENYLLHENCMLKKEIAMLK 1059
 Db 661 EEOHKELEVKQOLEQALRIQDIELKSVEINLVSHNENYLLHENCMLKKEIAMLK 720

Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
 Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 780

Qy 1120 TMLTSKLEKQDKLEILAEIESHPRLASAVODHDIQVTSRKSQBPFAHIAGDACLQRM 1179
 Db 781 TMLTSKLEKQDKLEILAEIESHPRLASAVODHDIQVTSRKSQBPFAHIAGDACLQRM 840

Qy 1180 NVDVSTIYNNNEVLHQPULSEAQRSKSLKINLVAGDALRENTLVSEHAQRDQRETQCOM 1239
 Db 841 NVDVSTIYNNNEVLHQPULSEAQRSKSLKINLVAGDALRENTLVSEHAQRDQRETQCOM 900

Qy 1240 KEAEHMYQNEODNVNKHTEQOESLDQKLFQLSKNMVLQOQLVHAHKADNKSKITIDIH 1299
 Db 901 KEAEHMYQNEODNVNKHTEQOESLDQKLFQLSKNMVLQOQLVHAHKADNKSKITIDIH 960

Qy 1300 FLERKMQHLLKKEKNEBIFNYYNNHLKNRIYQYEKEKATENS 1341
 Db 961 FLERKMQHLLKKEKNEBIFNYYNNHLKNRIYQYEKEKATENS 1002

RESULT 9
 ABG78918
 ID ABG78918 standard; protein; 1002 AA.
 XX AC ABG78918;
 XX DT 15-NOV-2002 (first entry)
 XX DE Human breast tumour polypeptide #10.
 XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
 XX OS Homo sapiens.
 XX FN US2002085998-A1.
 XX PD 04-JUL-2002.
 XX PF 13-APR-2001; 2001US-00834759.
 XX PR 28-DEC-1998; 98US-00222575.
 XX PR 02-APR-1999; 99US-00285480.
 XX PR 23-JUN-1999; 99US-00339338.
 XX PR 02-SEP-1999; 99US-00389681.
 XX PR 03-NOV-1999; 99US-00433826.
 XX PR 17-APR-2000; 2000US-00551621.
 XX PR 08-JUN-2000; 2000US-00590751.

CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
XX Sequence 1002 AA;

Query Match 74.1%; Score 5130; DB 6; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDORFPSESKQEDDEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDORFPSESKQEDDEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 120
Qy 460 PPKPSAFKPAIEMQNSVPNKAPELKNQTLRADPMFPPEKQKDYENSWDSLSLCEV 519
Db 121 PPKPSAFKPAIEMQNSVPNKAPELKNQTLRADPMFPPEKQKDYENSWDSLSLCEV 180
Qy 520 SKOVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQTFKABP 579
Db 181 SKOVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQTFKABP 240
Qy 580 PGKSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWDSLSLCEV 639
Db 241 PGKSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWDSLSLCEV 300
Qy 640 QKDVCPLKAAHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQTFKABP 699
Db 301 QKDVCPLKAAHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQTFKABP 360
Qy 700 EKPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWDSLSLCEV 759
Db 361 EKPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWDSLSLCEV 420
Qy 760 KDVCPLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQTFKABP 819
Db 421 KDVCPLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQTFKABP 480
Qy 820 KPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWDSLSLCEV 879
Db 481 KPSAFEPATEMOKSVNKALELKNQETWRADELTPSESKQKDYENSWDSLSLCEV 540
Qy 880 DVCVPKATHQKEMDKISGLKEDSTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 939
Db 541 DVCVPKATHQKEMDKISGLKEDSTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKK 600
Qy 940 FCVLKKLSEAKEIKSQLENQKWEQELCSVRLTLNQEEKRRNADILNEKIEELGRI 999
Db 601 FCVLKKLSEAKEIKSQLENQKWEQELCSVRLTLNQEEKRRNADILNEKIEELGRI 660
Qy 1000 EQHRKELEVKQLEQALRIODIELKSVESNLNOVSHNENYLLHNCMLKKEIAMLK 1059
Db 661 EQHRKELEVKQLEQALRIODIELKSVESNLNOVSHNENYLLHNCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 780
Qy 1120 TMLTSKLEKQDKEILEAIEESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQKRM 1179
Db 781 TMLTSKLEKQDKEILEAIEESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQKRM 840
Qy 1180 NVDVSSITLYNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCOM 1239
Db 841 NVDVSSITLYNEVLHQPSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRETQCOM 900
Qy 1240 KEABHYQNEQDNVNHKTEQOESLDOKLFQKSKNMWLQOQLVHAHKADNKSKITIDIH 1299

Db 901 KEABHYQNEQDNVNHKTEQOESLDOKLFQKSKNMWLQOQLVHAHKADNKSKITIDIH 960
Qy 1300 FLERKMQHLLKEKNEEIFFNNHLLKNRIYQVEKEKAETENS 1341
Db 961 FLERKMQHLLKEKNEEIFFNNHLLKNRIYQVEKEKAETENS 1002

RESULT 11
ADL93137
ID ADL93137 standard; protein; 1002 AA.
XX
AC ADL93137;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human breast cancer-associated polypeptide #10.
XX
KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX
OS Homo sapiens.
XX
PN US2003166022-A1.
XX
PD 04-SEP-2003.
XX
PF 15-APR-2002; 2002US-00124805.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH;
XX
XX WPI; 2003-874918/81.
DR N-PSDB; ADL93136.
XX
XX An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX
XX Example 1; SEQ ID NO 475; 294pp; English.
XX
XX The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a
CC cancer in a patient and treating a cancer in a patient. The present
CC sequence represents the amino acid sequence of a human breast cancer-
CC associated polypeptide.
XX
SQ Sequence 1002 AA;

Query Match 74.1%; Score 5130; DB 7; Length 1002;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACPTK 399
Db 1 MSPAKETSEKFTWAAGPRKIAWEKKEPTVKTGCVARVTSNKTIVLEKGRSKMIACPTK 60
Qy 400 ESSTKASANDORFPSESKQEDDEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 459
Db 61 ESSTKASANDORFPSESKQEDDEYSCDSRSLFPSSAKIQVCIPESIIYQKWEINREVEE 120

QY 460 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 519
 Db 121 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 180
 QY 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 579
 Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 240
 QY 580 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 639
 Db 241 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 300
 QY 640 QKDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 699
 Db 301 QKDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 360
 QY 700 EKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 759
 Db 361 EKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 420
 QY 760 KDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 819
 Db 421 KDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 480
 QY 820 KPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 879
 Db 481 KPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 540
 QY 880 DVCVPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 939
 Db 541 DVCVPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 600
 QY 940 FCVLKKLSEAKEIKSQLENQKWEQELCSVRLTLNOEEKRRNADILNEKIREELGRI 999
 Db 601 FCVLKKLSEAKEIKSQLENQKWEQELCSVRLTLNOEEKRRNADILNEKIREELGRI 660
 QY 1000 BEQHRKELEVKQLEQALRIQDIELKVESNLNQVSHTHENYLLHNCMLKKEIAMLK 1059
 Db 661 BEQHRKELEVKQLEQALRIQDIELKVESNLNQVSHTHENYLLHNCMLKKEIAMLK 720
 QY 1060 LEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKLKEESITKRASQYSGQLKVLIAEN 1119
 Db 721 LEIATLKHQYQEKENYFEDIKILKEKNAELQMTLKLKEESITKRASQYSGQLKVLIAEN 780
 QY 1120 TMLTSKLKEQDKEILEAIESHHPRASAVQDHDQIVTSRKSPQAFHAGDACLQRM 1179
 Db 781 TMLTSKLKEQDKEILEAIESHHPRASAVQDHDQIVTSRKSPQAFHAGDACLQRM 840
 QY 1180 NVDVSTIYNNEVLHQPSEAKQSKSLKINLYAGDALRENTLVSEHAQRDQTOCOM 1239
 Db 841 NVDVSTIYNNEVLHQPSEAKQSKSLKINLYAGDALRENTLVSEHAQRDQTOCOM 900
 QY 1240 KEAEMHYQNEQNVNKHTEQESLQKLQFQLSKNMLQQLVHAHKADNKSKITIDIH 1299
 Db 901 KEAEMHYQNEQNVNKHTEQESLQKLQFQLSKNMLQQLVHAHKADNKSKITIDIH 960
 QY 1300 FLERKQVHLLKKEKEEIPFNHNLKNRIYQYKEKAETENS 1341
 Db 961 FLERKQVHLLKKEKEEIPFNHNLKNRIYQYKEKAETENS 1002

RESULT 12
 ADE44427
 ID ADE44427 standard; protein; 1002 AA.
 XX
 AC ADE44427;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human breast cancer protein #10.
 XX human; breast tumour; cancer; vaccine; T cell stimulator;

T cell expander.
 Homo sapiens.
 US2003104366-A1.
 05-JUN-2003.
 17-APR-2000; 2000US-00551621.
 28-DEC-1998; 98US-00222575.
 02-APR-1999; 99US-00285480.
 23-JUN-1999; 99US-00339338.
 02-SEP-1999; 99US-00389681.
 03-NOV-1999; 99US-00433826.
 (JIAN)/ JIANG Y.
 (DILL)/ DILLON D C.
 (MITC)/ MITCHAM J L.
 (XUJJ)/ XU J.
 (HARL)/ HARLOCKER S L.
 Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
 WPI; 2004-020270/02.
 N-PSDB; ADE44426.
 Novel isolated polypeptide comprising immunogenic portion of breast tumor protein or its variant, useful for formulating vaccines for inhibiting cancer development in a patient.
 Example 1; SEQ ID NO 475; 217pp; English.
 The invention relates to an isolated polypeptide comprising at least an immunogenic portion of a breast tumour protein. The polynucleotide, its fusion protein or the polynucleotide encoding it, a vaccine comprising the fusion protein or the polynucleotide encoding it, an isolated T cell, a population comprising T cells specific for a breast tumour protein, and a method for removing tumour cells from a biological sample is useful for inhibiting the development of a cancer in a patient. The polypeptide is useful for stimulating and/or expanding T cells specific for a breast tumour protein. Stimulating and/or expanding T cells specific for a breast tumour protein is useful for inhibiting the development of a cancer in a patient. The method additionally involves the step of cloning at least one proliferated cell and then administering the cloned T cells to the patient. The present sequence represents a breast cancer protein.
 Sequence 1002 AA;
 Query Match 74.1%; Score 5130; DB 8; Length 1002;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 340 MSPAKETSEKPTWAAGRPRIAEKKEPTVKTGCVARVTSNKTIVLEKSGSKMIACPTK 399
 Db 1 MSPAKETSEKPTWAAGRPRIAEKKEPTVKTGCVARVTSNKTIVLEKSGSKMIACPTK 60
 QY 400 ESSTKASANDQFPSEKQEEDEYSCDSRSLFPSSAKIOVCIPESIIYQKWEINREVEE 459
 Db 61 ESSTKASANDQFPSEKQEEDEYSCDSRSLFPSSAKIOVCIPESIIYQKWEINREVEE 120
 QY 460 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 519
 Db 121 PPKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 180
 QY 520 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 579
 Db 181 SQKDVCLPKATHQKEIDKINGKLEBSPNKGGLKATCGMKVSIPTKALELMDQMTFKAEP 240
 QY 580 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 639
 Db 241 PGKPSAFKPAIEMQNSVPNKAFAELKNEQTLRADMPPEPSKQKDYENSWDSLSLCTV 300


```
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 699
Db 301 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 360
Qy 700 EKPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 759
Db 361 EKPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 420
Qy 760 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 819
Db 421 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 480
Qy 820 KPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 879
Db 481 KPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 540
Qy 880 DVCVPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 939
Db 541 DVCVPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 600
Qy 940 FCVLKCKLSEAKEIKSQLENQKWEQELCSVRLTLNOEERKRNADILNEKIREELGRI 999
Db 601 FCVLKCKLSEAKEIKSQLENQKWEQELCSVRLTLNOEERKRNADILNEKIREELGRI 660
Qy 1000 EEOHKELEVKQOQLEQALRIQDIELKSYESNLNOVSHTHENYLLHENCMLKKEIAMLK 1059
Db 661 EEOHKELEVKQOQLEQALRIQDIELKSYESNLNOVSHTHENYLLHENCMLKKEIAMLK 720
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTRKASQYSGQLKVLIAEN 1119
Db 721 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTRKASQYSGQLKVLIAEN 780
Qy 1120 TMLTSKLEKQDKETLEAEIESHPRLASAVQDDHDIIVTSKSOEPAFHIAAGACLQKRM 1179
Db 781 TMLTSKLEKQDKETLEAEIESHPRLASAVQDDHDIIVTSKSOEPAFHIAAGACLQKRM 840
Qy 1180 NVDVSSITYNNEVLHQLPSEAKRSKSLKINLNTAGDALRENTLVSEHAQRDQRETQCM 1239
Db 841 NVDVSSITYNNEVLHQLPSEAKRSKSLKINLNTAGDALRENTLVSEHAQRDQRETQCM 900
Qy 1240 KEAHEMYQNEQDNVNHKTEQESLDQKLFLQSKNMWLOQQLVHAHKADNKSITIDIH 1299
Db 901 KEAHEMYQNEQDNVNHKTEQESLDQKLFLQSKNMWLOQQLVHAHKADNKSITIDIH 960
Qy 1300 FLERQWQHLLKXNEEETFNYNHLLKNRIYQYEKEKAETENS 1341
Db 961 FLERQWQHLLKXNEEETFNYNHLLKNRIYQYEKEKAETENS 1002

RESULT 13
AAU33357
ID AAU33357 standard; protein; 1095 AA.
AC AAU33357;
XX
XX
DT 18-DEC-2001 (first entry)
XX
DE Human breast cancer protein B726P fusion protein #1.
XX
XX Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX gene therapy; immunogen.
XX
XX Homo sapiens.
XX
XX WO200179286-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012164.
XX
XX 17-APR-2000; 2000US-00551621.
XX
XX 08-JUN-2000; 2000US-005590751.
```

```
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
DR N-PSDB; AAS47421.
DR
```

Breast Tumor Proteins and nucleic acids useful for the prevention, diagnosis and treatment of breast cancer.

Claim 22; Page 292-295; 297pp; English.

The invention relates to isolated breast tumour proteins and nucleic acids that encode them, including immunogenic fragments of the proteins. Also included are expression vectors expressing the proteins, transformed cells and antibodies raised against the proteins or an antigen presenting cell expressing the protein. The proteins and nucleic acids may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate breast tumour protein expression, i.e. breast tumours and breast cancer e.g. by gene therapy. The nucleic acids and their complements may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The proteins, nucleic acids and antibodies may be used in assays to identify modulators (e.g. antagonists) of breast tumour protein expression and activity. The antibodies and antagonists may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the proteins in samples (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immunopurification diagnostic techniques. The present sequence is a breast tumour protein encoded by a cDNA from a breast tumour cDNA library isolated by subtractive hybridisation against a normal breast cDNA library

Sequence 1095 AA;

Query Match 74.1%; Score 5130; DB 4; Length 1095;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy 340 MSPAKETSEKFTWAAGRPRIAKKEKTPVKTGCVARVTSNKTKEGSKMIACPTK 399
Db 94 MSPAKETSEKFTWAAGRPRIAKKEKTPVKTGCVARVTSNKTKEGSKMIACPTK 153
Qy 400 ESSTKASANDQRPFPSEKQDEEYSCDSRSLFPSSAKIQVICPESIIYQKMEINREVEE 459
Db 154 ESSTKASANDQRPFPSEKQDEEYSCDSRSLFPSSAKIQVICPESIIYQKMEINREVEE 213
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADEILPSESKQKDYESSWDSSESLCETV 519
Db 214 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADEILPSESKQKDYESSWDSSESLCETV 273
Qy 520 SOKDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 579
Db 274 SOKDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 333
Qy 580 PGKPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVS 639
Db 334 PGKPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVS 393
Qy 640 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 699
Db 394 QKDVCLPKAAHQKEIDKINGKLESPVKDGLLKANCCKMVKSIPTKALELMDMOTFKAEPP 453
Qy 700 EKPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 759
Db 454 EKPSAFPAIEMQKSVPNKALELNEQTLRADEILPSESKQKDYESSWDSSESLCETVSQ 513
Qy 760 KDVCLPKATHQKEIDKINGKLESPNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPP 819
```

Db 514 KDVCLPKAXHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAEPPPE 573
Qy 820 KPSAFPAIEMQKSVNPKALELKNQOTLRADQMPFSESKQKVEENSWDSLSLRETVSQK 879
Db 574 KPSAFPAIEMQKSVNPKALELKNQOTLRADQMPFSESKQKVEENSWDSLSLRETVSQK 633
Qy 880 DVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 939
Db 634 DVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMKKK 693
Qy 940 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIRELGRI 999
Db 694 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIRELGRI 753
Qy 1000 EEOHKELEVKQOOLFOALRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLK 1059
Db 754 EEOHKELEVKQOOLFOALRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTRASQYSGQLKVLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTRASQYSGQLKVLIAEN 873
Qy 1120 TMLTSKLKEKQDKEILEAEIESHPRLASAVQDHDQIVTSRKSQBPAPHIAGDAQLQRM 1179
Db 874 TMLTSKLKEKQDKEILEAEIESHPRLASAVQDHDQIVTSRKSQBPAPHIAGDAQLQRM 933
Qy 1180 NVDVSSSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRTQCOM 1239
Db 934 NVDVSSSTIYNNEVLHQPILSEAKRSKSLKINLYAGDALRENTLVSEHAQRDQRTQCOM 993
Qy 1240 KEAEHMYQNEQDNVNHKTEQOESLDQKLFQLSKNMWLQQQLVHAHKADNKSKITIDIH 1299
Db 994 KEAEHMYQNEQDNVNHKTEQOESLDQKLFQLSKNMWLQQQLVHAHKADNKSKITIDIH 1053
Qy 1300 FLERQMOHLLKEKNEEFNYYNHLKNRIYQVEKEKAETENS 1341
Db 1054 FLERQMOHLLKEKNEEFNYYNHLKNRIYQVEKEKAETENS 1095

RESULT 14

ABG78924

ID ABG78924 standard; protein; 1095 AA.

XX

AC ABG78924;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human breast tumour polypeptide #15.

XX

KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.

XX

OS Homo sapiens.

XX

PN US2002085998-A1.

XX

PD 04-JUL-2002.

XX

PF 13-APR-2001; 2001US-00834759.

XX

PR 28-DEC-1998; 98US-00222575.

XX

PR 02-APR-1999; 99US-00285480.

XX

PR 23-JUN-1999; 99US-00339338.

XX

PR 02-SEP-1999; 99US-00389681.

XX

PR 03-NOV-1999; 99US-00433826.

XX

PR 17-APR-2000; 2000US-00551621.

XX

PR 08-JUN-2000; 2000US-00590751.

XX

PR 22-JUN-2000; 2000US-00604287.

XX

PR 20-JUL-2000; 2000US-00620405.

XX

PA (CORI-) CORIXA CORP.

XX

PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;

PI Henderson RA;

XX

DR WPI; 2002-635657/68.

XX N-PSDB; ABS64022.

XX

Novel breast cancer polynucleotides and polypeptides encoded by the polynucleotides, useful for detecting the presence of breast cancer in a patient, and in pharmaceutical compositions, for treating breast cancer.

XX

PS Disclosure; Page 223-225; 247pp; English.

XX

The invention relates to an isolated breast tumour polynucleotide and the polypeptide it encodes. The polynucleotide and polypeptide are useful for detecting the presence of breast cancer in a patient, and in pharmaceutical compositions for treating breast cancer. The sequences are useful for stimulating an immune response in a patient and can therefore be used in production of vaccines. The sequences are also useful for detecting the presence of a cancer in a patient, by obtaining a biological sample from the patient, contacting the biological sample with a composition of the invention and detecting the amount of polynucleotide that hybridizes to the sample. This sequence represents a human breast tumour polypeptide of the invention

XX

SQ Sequence 1095 AA;

XX

Query Match 74.1%; Score 5130; DB 5; Length 1095;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy

340 MSPAKETSEKPTWAAGRPRIAEKKEPTVTCGVARVTSNKTQVLEKSGSKMIACPTK 399

Db

94 MSPAKETSEKPTWAAGRPRIAEKKEPTVTCGVARVTSNKTQVLEKSGSKMIACPTK 153

Qy

400 ESSSTKASANDORPFSESKQDEEYSCDSRLSFESSAKIQVCIPESIQKWEINREVEE 459

Db

154 ESSSTKASANDORPFSESKQDEEYSCDSRLSFESSAKIQVCIPESIQKWEINREVEE 213

Qy

460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMPFPEKSKQDYENSWDSLSLCTV 519

Db

214 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMPFPEKSKQDYENSWDSLSLCTV 273

Qy

520 SQDVCLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDQMTFKAE 579

Db

274 SQDVCLPKATHQKEIDKINGKLEESPNKDGLLKATCGMKVSIPTKALELMDQMTFKAE 333

Qy

580 PGKPSAPAPATEMOKSVNPKALELKNQOTLRADQMPFSESKQDYENSWDSLSLCTVS 639

Db

334 PGKPSAPAPATEMOKSVNPKALELKNQOTLRADQMPFSESKQDYENSWDSLSLCTVS 393

Qy

640 QKDVCPLKAAHQKEIDKINGKLEESPVKDGILLKANCQMKVSIPTKALELMDQMTFKAE 699

Db

394 QKDVCPLKAAHQKEIDKINGKLEESPVKDGILLKANCQMKVSIPTKALELMDQMTFKAE 453

Qy

700 EKPSAFPAIEMQKSVNPKALELKNQOTLRADQMPFSESKQDYENSWDSLSLCTVSQ 759

Db

454 EKPSAFPAIEMQKSVNPKALELKNQOTLRADQMPFSESKQDYENSWDSLSLCTVSQ 513

Qy

760 KDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAE 819

Db

514 KDVCLPKATHQKEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDQMTFKAE 573

Qy

820 KPSAFPAIEMQKSVNPKALELKNQOTLRADQMPFSESKQDYENSWDSLSLRETVSQK 879

Db

574 KPSAFPAIEMQKSVNPKALELKNQOTLRADQMPFSESKQDYENSWDSLSLRETVSQK 633

Qy

880 DVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMK 939

Db

634 DVCVPKATHQKEMDKISGKLESDTSLSKILDTVHSCERARELQKDHCEQRTGKMEQMK 693

Qy

940 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIRELGRI 999

Db

694 FCVLKKLSEAKEIKSQLENQKVWEQELCSVRLTLNOEEKRRNADILNEKIRELGRI 753

Qy

1000 EEOHKELEVKQOOLFOALRIODIELKSVESNLNQVSHTHENENYLLHNCMLKKEIAMLK 1059

Db 754 EEQHRKEVKKQLEQALRIQDIELKSVESNLQVSHTHENYLLHENCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 873
Qy 1120 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 1179
Db 874 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 933
Qy 1180 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 934 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 993
Qy 1240 KEAEMHTQNEODNVNKHTEQOESLDQKLFQSQKNMMLQOQLVHAHKADNKSITIDIH 1299
Db 994 KEAEMHTQNEODNVNKHTEQOESLDQKLFQSQKNMMLQOQLVHAHKADNKSITIDIH 1053
Qy 1300 FLERKMOHLLKKEKNEEIFNYYNNHLKNRIYOEKEKAETENS 1341
Db 1054 FLERKMOHLLKKEKNEEIFNYYNNHLKNRIYOEKEKAETENS 1095

RESULT 15

ABJ37747
ID ABJ37747 standard; protein; 1095 AA.

AC ABJ37747;

XX 15-MAY-2003 (first entry)

DE Human tumour-related protein - SEQ ID No 493.

XX Human: vaccine; gene therapy; T cell stimulation; T cell expansion;
XX tumour; breast cancer; cancer; immune response stimulation.

XX Homo sapiens.

XX WO200283956-A1.

XX 24-OCT-2002.

XX 15-APR-2002; 2002WO-US012378.

XX 13-APR-2001; 2001US-00834759.

XX 07-DEC-2001; 2001US-00007805.

XX 13-FEB-2002; 2002US-00076822.

XX (CORI-) CORIXA CORP.

XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;

XX Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;

XX Vedvick TS, McNeill PD, Durham M;

XX WPI; 2003-103376/09.

XX New polypeptide and polynucleotide useful for stimulating and/or

XX expanding T cells specific for a tumor protein and treating breast

XX cancer.

XX Disclosure; Page 310-312; 375pp; English.

XX The invention comprises a method of stimulating and/or expanding T cells

XX specific for a tumour protein. The invention further comprises human

XX nucleic acids and proteins that are associated with tumours (e.g. breast

XX cancer). The method and sequences of the invention are useful for

XX stimulating and/or expanding T cells specific for a tumour protein,

XX detecting the presence of cancer, stimulating an immune response in a

XX patient and treating breast cancer. The present amino acid sequence

XX represents a human tumour-related protein

XX Sequence 1095 AA;

Query Match 74.1%; Score 5130; DB 6; Length 1095;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 996; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 340 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMTACPTK 399
Db 94 MSPAKETSEKFTWAAKGRPRKIAWEKKEPTVKTCGVARVTSNKTQVLEKGRSKMTACPTK 153
Qy 400 ESSTKASANDORPPSESKQDEEYSCDSRSLFESSAKIQVCIPESIIQKWEINREVEE 459
Db 154 ESSTKASANDORPPSESKQDEEYSCDSRSLFESSAKIQVCIPESIIQKWEINREVEE 213
Qy 460 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMPFPPEKSKQDYEENSWSLCEVT 519
Db 214 PPKPSAFKPAIEMQNSVPNKAFELKNEQTLRADPMPFPPEKSKQDYEENSWSLCEVT 273
Qy 520 SQKDVCLPKATHQKEIDKINGKLEESPNGDGLLKATCGMKVSIPTKALELMDQTFKABPP 579
Db 274 SQKDVCLPKATHQKEIDKINGKLEESPNGDGLLKATCGMKVSIPTKALELMDQTFKABPP 333
Qy 580 PGKPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPPEKSKQDYEENSWSLCEVT 639
Db 334 PGKPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPPEKSKQDYEENSWSLCEVT 393
Qy 640 QKDVCLPKAAHQKEIDKINGKLEESPNGDGLLKATCGMKVSIPTKALELMDQTFKABPP 699
Db 394 QKDVCLPKAAHQKEIDKINGKLEESPNGDGLLKATCGMKVSIPTKALELMDQTFKABPP 453
Qy 700 EKPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPPEKSKQDYEENSWSLCEVT 759
Db 454 EKPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPPEKSKQDYEENSWSLCEVT 513
Qy 760 KDVCLPKATHQKEIDKINGKLEESPNDGDLKAPCRMVSIPTKALELMDQTFKABPP 819
Db 514 KDVCLPKATHQKEIDKINGKLEESPNDGDLKAPCRMVSIPTKALELMDQTFKABPP 573
Qy 820 KPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPPEKSKQDYEENSWSLCEVT 879
Db 574 KPSAFEPATEMOKSVPNKALELKNQETLRADPMPFPPEKSKQDYEENSWSLCEVT 633
Qy 880 DVCVPKATHQKEMDKISGKLEDDSTLSKILDTVHSCERARELQKDHQCHQRTGKMBQMKK 939
Db 634 DVCVPKATHQKEMDKISGKLEDDSTLSKILDTVHSCERARELQKDHQCHQRTGKMBQMKK 693
Qy 940 FCVLKKKLSAKEITKSOLEKQVKEQELCSVRITLQOEERKRRNADILNEKIREELGRI 999
Db 694 FCVLKKKLSAKEITKSOLEKQVKEQELCSVRITLQOEERKRRNADILNEKIREELGRI 753
Qy 1000 EEQHRKEVKKQLEQALRIQDIELKSVESNLQVSHTHENYLLHENCMLKKEIAMLK 1059
Db 754 EEQHRKEVKKQLEQALRIQDIELKSVESNLQVSHTHENYLLHENCMLKKEIAMLK 813
Qy 1060 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 1119
Db 814 LEIATLKHQYQEKENKYPEDIKILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAEN 873
Qy 1120 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 1179
Db 874 TMLTSKLKQKQDKEILEAEIESHPRLASAVODHDQIVTSRKSQEPAPHIAGDACLQRM 933
Qy 1180 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 1239
Db 934 NVDVSSIIYNNVHLQPLSEARQSKSLKINLYAGDALRENTLVSEHAQDQRETQCOM 993
Qy 1240 KEAEMHTQNEODNVNKHTEQOESLDQKLFQSQKNMMLQOQLVHAHKADNKSITIDIH 1299
Db 994 KEAEMHTQNEODNVNKHTEQOESLDQKLFQSQKNMMLQOQLVHAHKADNKSITIDIH 1053
Qy 1300 FLERKMOHLLKKEKNEEIFNYYNNHLKNRIYOEKEKAETENS 1341
Db 1054 FLERKMOHLLKKEKNEEIFNYYNNHLKNRIYOEKEKAETENS 1095

Search completed: February 19, 2005, 02:04:43
Job time : 151.087 secs
